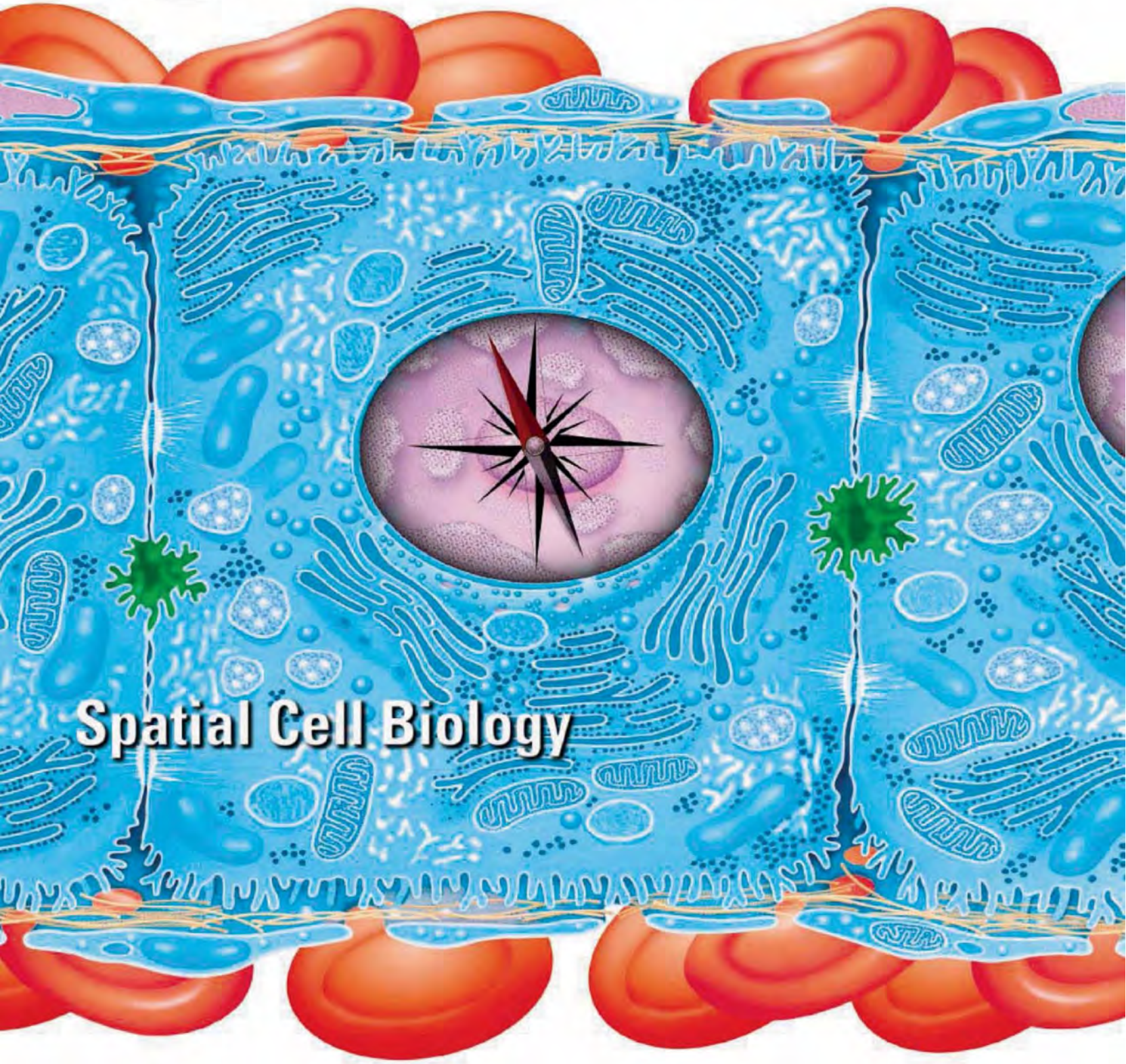
27 November 2009 | $10

Science



Spatial Cell Biology

AAAS

electrofocused.

Focused on your complete electrophoresis solutions.

With more than 50 years of expertise in electrophoresis, Bio-Rad is focused on providing complete solutions for reliable, reproducible results for all of your electrophoretic separation needs.

Backed by award-winning technical support, our Mini-PROTEAN, Criterion and PROTEAN electrophoresis platforms cover your full range of applications. From easy-to-use precast and handcast gels to fully integrated blotting systems, we have a solution for you.

Gel with us by visiting **www.bio-rad.com/ad/electrofocused**.

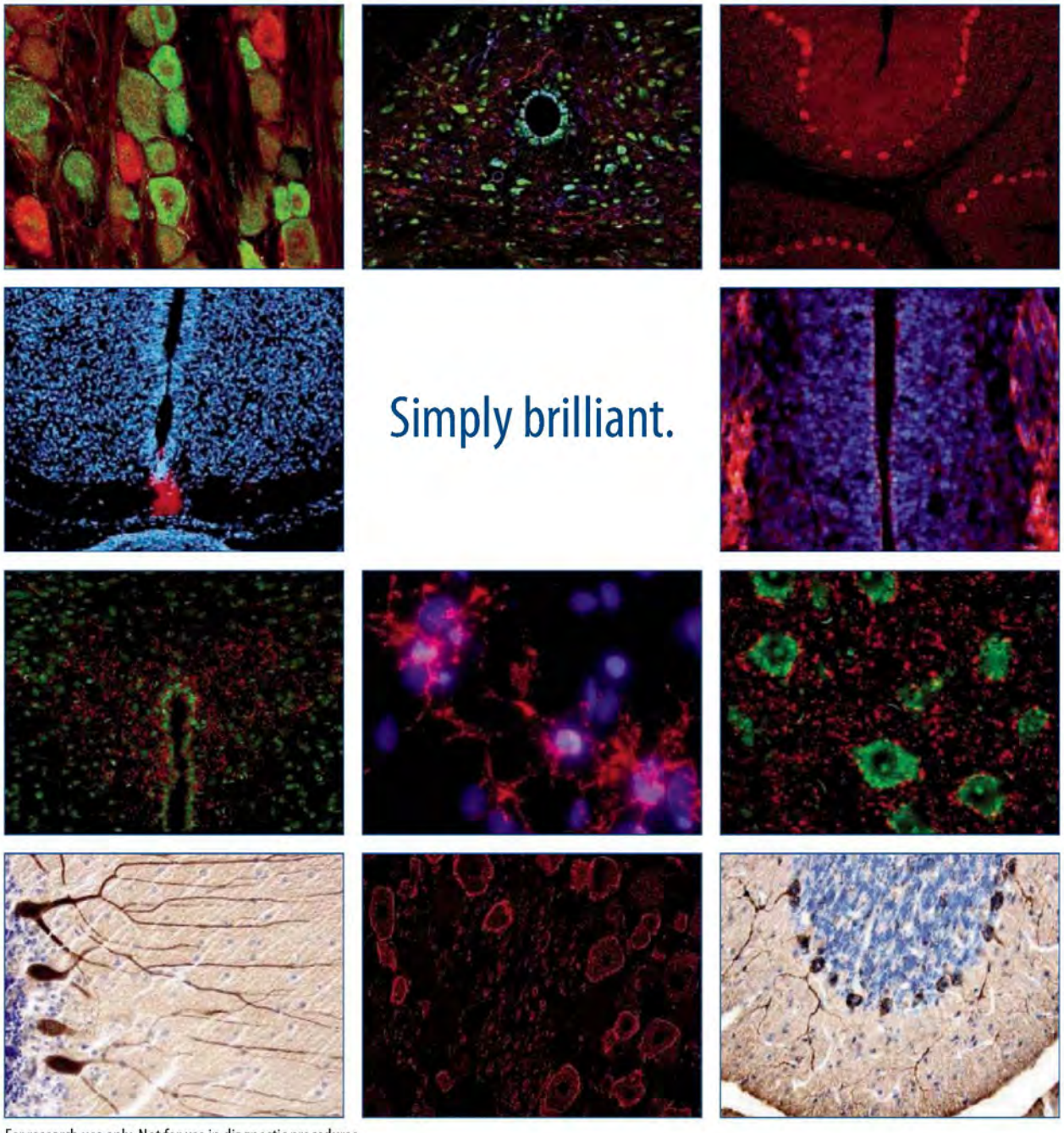
Research. Together.

To find your local sales office, visit **www.bio-rad.com/contact/**
In the U.S., call toll free at 1-800-4BIORAD (1-800-424-6723)

Visit us at www.bio-rad.com

BIO-RAD

R&D Systems
Antibodies for Neuroscience Research



Simply brilliant.

For research use only. Not for use in diagnostic procedures.

For more information visit our website at **www.RnDSystems.com/go/Neuroscience**

R&D Systems Tools for Cell Biology Research™

USA & Canada **R&D Systems, Inc.** Tel: (800) 343-7475 info@RnDSystems.com

Europe **R&D Systems Europe, Ltd.** Tel: +44 (0)1235 529449 info@RnDSystems.co.uk

Selection expanding weekly—visit **www.RnDSystems.com/go/request** to sign up for weekly new product updates.

R&D SYSTEMS®

GE Healthcare
Life Sciences

Inspired Again

Who better to draw inspiration for the new ÄKTA™ avant system than from customers using the 30,000 ÄKTA systems already in use around the world? Well, you spoke and we listened. The new ÄKTA avant system for process development is faster — enabling quicker insights. It minimizes the chance of error, even while working at higher speeds. And it allows for direct, reliable scalability. At GE Healthcare, our focus is on helping scientists achieve even more, faster. It's a commitment we have in our genes. And all this is backed by the service, support, and investment in the future that being part of GE can bring.

Want to know more? Why not talk with us today. Visit **www.gelifesciences.com/aktaavant**

I **ÄKTA** I Amersham I Biacore I IN Cell Analyzer I Whatman I GE Service I

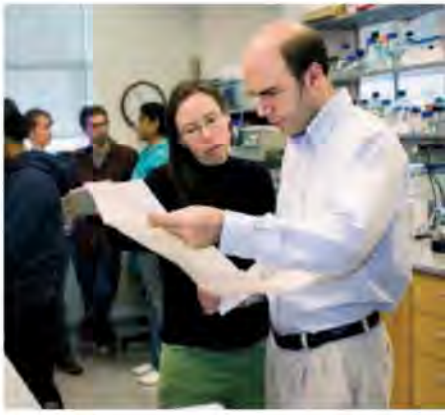
The New ÄKTA avant

imagination at work

First published September 2009
GE Healthcare Bio-Sciences AB, Björkgatan 30, 751 84 Uppsala, Sweden
GE12-09

SPECIAL SECTION

Spatial Cell Biology



page 1179

CONTENTS continued >>

COVER

Liver cells possess a well-defined morphology and interact in time and space with their neighbors, with the extracellular matrix, and with the bloodstream. A special section starting on page 1205 describes some of the basic tenets of spatial cell biology.

Image: Frank Geisler/Alamy (liver cells); iStockphoto.com (compass rose)

DEPARTMENTS

Analyzing genetic differences

Genotyping sample and assay technologies by QIAGEN

Rely on trusted automated and manual workflow solutions for:

- **Sample collection and stabilization**
- **Genomic DNA purification, DNA storage, and whole genome amplification**
- **PCR amplification and automated QIAxcel® fragment analysis**
- **HRM® and Pyrosequencing® detection**

Making improvements in life possible — www.qiagen.com

QIAGEN®

ADNAGenotyp0909051WW

Sample & Assay Technologies

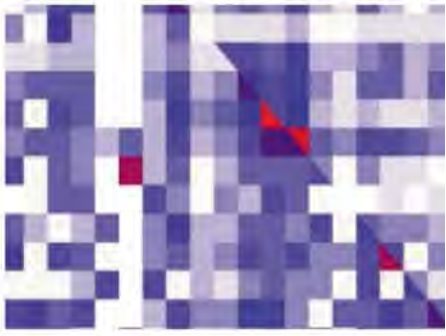
BREVIA

RESEARCH ARTICLES

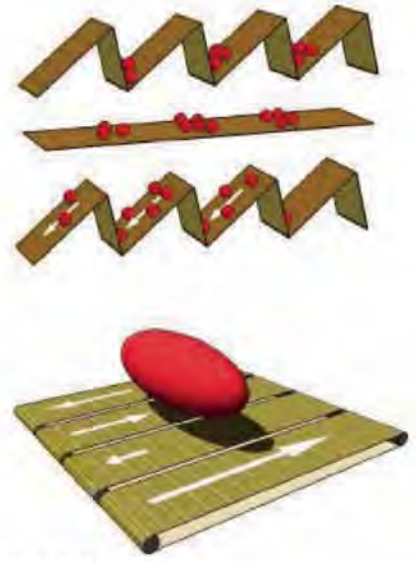
REPORTS

CONTENTS continued >>

pages 1194 & 1244



pages 1200, 1235, 1263, & 1268



page 1241

ampliflexible.

Flexible, high-throughput qPCR.

Bio-Rad's CFX384™ real-time PCR system is the best solution for your high-throughput qPCR needs. It can be adapted to your unique workflow to accelerate your discoveries with the unparalleled performance and reliability you know and trust from Bio-Rad.

- Get results faster with the streamlined experiment setup and powerful data analysis tools
- Save samples and reagents with low volume 4-target multiplex reactions; achieve optimal results with pre-developed assays
- Maximize your productivity even more with the CFX automation system, with walk away, worry-free automated operation

Bring our 20 years of experience in PCR to your discovery. Visit **www.bio-rad.com/ad/amplifamily/** or contact your Bio-Rad sales representative to learn more.

Research. Together.

To find your local sales office, visit **www.bio-rad.com/contact/**
In the U.S., call toll free at 1-800-4BIORAD (1-800-424-6723)

Visit us at www.bio-rad.com

BIO-RAD

SCIENCEONLINE

SCIENCEXPRESS
www.sciencexpress.org

Modulated High-Energy Gamma-Ray Emission from the Microquasar Cygnus X-3
The Fermi LAT Collaboration
Gamma-ray emission from the jet of an accreting binary star system is correlated with the jet's radio emission.
10.1126/science.1182174

Ligand-Enabled Reactivity and Selectivity in a Synthetically Versatile Aryl C–H Olefination
D.-H. Wang et al.
A palladium-based catalyst eliminates the need for halogenated compounds for the formation of carbon-carbon bonds.
10.1126/science.1182512

On the Origin of Species by Natural and Sexual Selection
G. S. van Doorn et al.
Modeling demonstrates how speciation occurs due to sexual selection.
10.1126/science.1181661
>> Science *Podcast*

Tetrathiomolybdate Inhibits Copper Trafficking Proteins Through Metal Cluster Formation
H. M. Alvarez et al.
Complex formation between a copper chaperone and a metallo-drug prevents copper transfer to target enzymes.
10.1126/science.1179907

SCIENCENOW
www.sciencenow.org
Highlights From Our Daily News Coverage

Want to Stop AIDS? Spend Big
A new model argues that a massive cash infusion could halt the spread of the disease.

Slideshow: Ancient Crocs With a Dog-Like Walk
Scientists find a strange collection of fossils in the African desert.

Birth of a (Magnetic) Heavyweight
Observations show how magnetic fields guide the emergence of a massive new star.

SCIENCESIGNALING
www.sciencesignaling.org
The Signal Transduction Knowledge Environment

RESEARCH ARTICLE: mTOR Regulation and Therapeutic Rejuvenation of Aging Hematopoietic Stem Cells
C. Chen et al.
Rapamycin reverses aging-related declines in hematopoietic stem cell function.

RESEARCH ARTICLE: Eukaryotic Protein Domains as Functional Units of Cellular Evolution
J. Jin et al.
Clustering proteins into groups on the basis of their domain compositions provides insight into protein evolution.

RESEARCH ARTICLE: Ca^{2+} Puffs Originate from Pre-Established Stable Clusters of Inositol Trisphosphate Receptors
I. F. Smith et al.
Localized calcium signals called Ca^{2+} puffs arise at pre-established clusters of IP_3Rs.

PERSPECTIVE: Maintaining Diplomatic Relations Between Mammals and Beneficial Microbial Communities
D. A. Hill and D. Artis
The adaptive immune system compensates if innate mechanisms fail to contain microbes in the mammalian intestine.

PODCAST
T. Pawson and A. M. VanHook
Bioinformatics analysis reveals how protein domain composition correlates with evolutionary change.

SCIENCECAREERS
www.sciencecareers.org/career_magazine
Free Career Resources for Scientists

Careers in Climate Change Research: Feature Index
E. Pain
Opportunities are expanding for natural scientists willing to tackle climate change.

Climate Science Broadens to Meet New Challenges
S. Carpenter
New programs help prepare scientists for the interdisciplinary nature of climate change research.

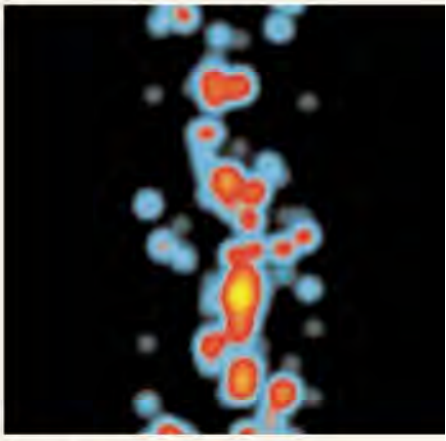
On-the-Ground Training for Climate Change Researchers
E. Pain
Climate change scientists need a unique blend of skills that often must be acquired informally.

SCIENCETRANSLATIONAL MEDICINE
www.sciencetranslationalmedicine.org
Integrating Medicine and Science

COMMENTARY: Use of Forensic Methods Under Exigent Circumstances Without Full Validation
S. E. Schutzer et al.
In emergency situations, forensic methods need preliminary validation to ensure accurate interpretation.

RESEARCH ARTICLE: Efficacy of Cetuximab in the Treatment of Ménétrier's Disease
W. H. Fiske et al.

PERSPECTIVE: Ménétrier's Disease Therapy: Rebooting Mucosal Signaling
S. C. Nalle et al.
A new study offers insight into the causes of a rare stomach disorder and describes a promising targeted therapeutic approach.



CREDIT: XUEYING XIE AND YONGYI MAO/SCHOOL OF INFORMATION TECHNOLOGY AND ENGINEERING (SITE)/UNIVERSITY OF OTTAWA, ONTARIO, CANADA

SCIENCESIGNALING
Visualizing the molecular environments of protein dynamics.

RESEARCH ARTICLE: In Situ Regulation of DC Subsets and T Cells Mediates Tumor Regression in Mice
O. A. Ali et al.
An implanted matrix elicits an immune response network that can eradicate established tumors in mice.

SCIENCEPODCAST
www.sciencemag.org/multimedia/podcast
Free Weekly Show
Download the 27 November *Science* Podcast to hear about speciation by sexual selection, a medieval climate anomaly, the fate of stimulus funding for science, and more.

ORIGINSBLOG
blogs.sciencemag.org/origins
A History of Beginnings

SCIENCEINSIDER
blogs.sciencemag.org/scienceinsider
Science Policy News and Analysis

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): $146 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $835; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $20.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

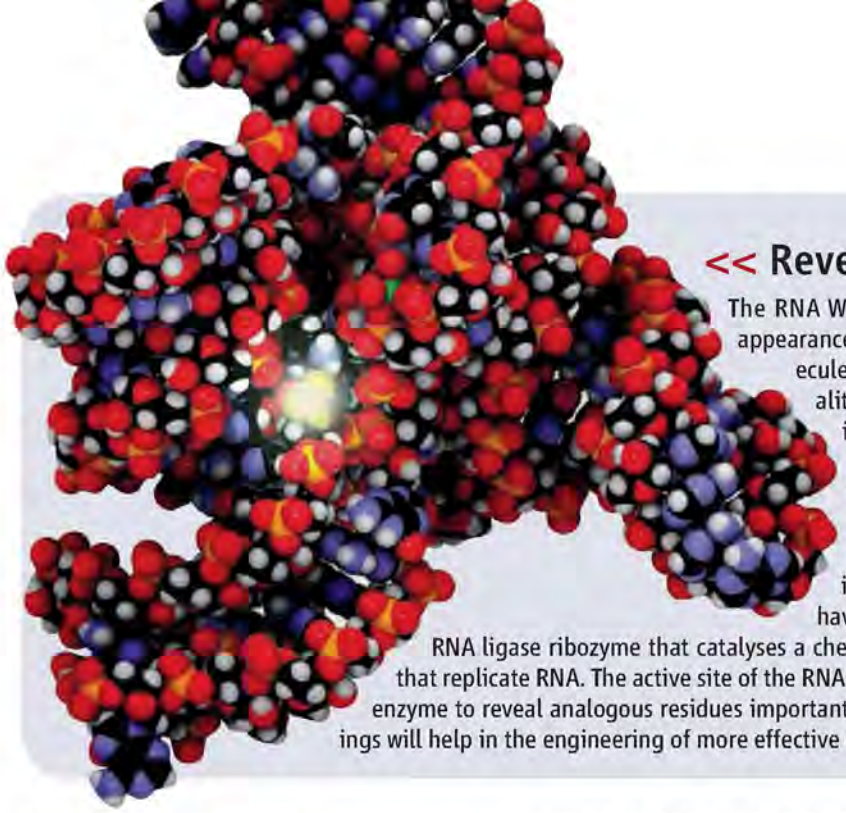
The Cure for Poor Resolution

MeltDoctor™ HRM Reagents

Superior resolution is the key to achieving higher accuracy and reproducibility in HRM analysis. MeltDoctor™ HRM reagents from Applied Biosystems provide best-in-class performance, allowing you to discriminate between wild-type and variant samples with greater resolution than competing mixes. MeltDoctor™ HRM reagents are optimized for a wide range of targets and variants, and provide a streamlined, closed-tube workflow that helps reduce the risk of sample contamination. Combined with Applied Biosystems world-class real-time PCR instruments and HRM Software, these reagents offer a complete workflow solution that enables you to obtain the clarity you need and results you can trust.

Improve the health of your HRM experiments:
www.appliedbiosystems.com/hrm

applied biosystems™



<< Revealing the RNA World?

The RNA World hypothesis posits that at an early step in the appearance of life, RNA acted both as an information storage molecule and as an enzyme—or ribozyme. Such dual functionality would allow for an RNA species that could replicate itself and thus seed the beginning of molecular evolution. The involvement of RNA in a number of fundamental cell biological processes, together with its ability, either naturally or through in vitro evolution, to catalyze a range of chemical reactions, provides some indirect support for this view. **Shechner *et al.*** (p. 1271) have now determined the structure of an in vitro–evolved RNA ligase ribozyme that catalyses a chemical reaction essentially identical to that of proteins that replicate RNA. The active site of the RNA ligase could be superimposed upon that of the protein enzyme to reveal analogous residues important for the catalytic joining of RNA moieties. These findings will help in the engineering of more effective ribozyme polymerases.

"Haploid Human"

Genetic screens can provide direct insight into biological processes that are poorly understood. **Carette *et al.*** (p. 1231) describe genetic screens using large-scale gene disruption in human cells haploid for all chromosomes except for chromosome 8. One screen was used to identify host factors essential for the activity of cytolethal distending toxin, a toxin found in several pathogenic bacteria. Another screen identified host gene products essential for infection with influenza, and an additional screen revealed genes required for the action of adenosine 5′-diphosphate (ADP)–ribosylating bacterial toxins. This loss-of-function genetic approach in mammalian cells will be widely applicable to study a variety of biological processes and cellular functions.

Simply *Mycoplasma*

The bacterium *Mycoplasma pneumoniae*, a human pathogen, has a genome of reduced size and is one of the simplest organisms that can reproduce outside of host cells. As such, it represents an excellent model organism in which to attempt a systems-level understanding of its biological organization. Now three papers provide a comprehensive and quantitative analysis of the proteome, the metabolic network, and the transcriptome of *M. pneumoniae* (see the Perspective by **Ochman and Raghavan**). Anticipating what might be possible in the future for more complex organisms, **Kühner *et al.*** (p. 1235) combine analysis of protein interactions by mass spectrometry with extensive structural information on *M. pneumoniae* proteins to reveal how proteins work together as molecular machines and map their organization within the cell by electron tomography. The manageable genome size of *M. pneumoniae* allowed **Yus *et al.*** (p. 1263) to map the metabolic network of the organism manually and validate it experimentally. Analysis of the network aided development of a minimal medium in which the bacterium could be cultured. Finally, **Güell *et al.*** (p. 1268) applied state-of-the-art sequencing techniques to reveal that this "simple" organism makes extensive use of noncoding RNAs and has exon- and intron-like structure within transcriptional operons that allows complex gene regulation resembling that of eukaryotes.

Crystal Growing Kit

For single crystals to remain intact, there is a limit to the size and number of defects that can be included before the underlying lattice is destroyed. Biological crystals, however, are known to include large macromolecules. **H. Li *et al.*** (p. 1244; see the Perspective by **Hollingsworth**) used electron tomography to study the crystallization of calcium carbonate inside an agarose gel, observing that the crystals physically entrapped the agarose macromolecules. To accommodate the curvature induced by the polymer chains, both high- and low-energy facets formed at the fiber-crystal interfaces. Thus, physical interactions alone may be sufficient for the incorporation of macromolecules in biological crystals and it may be possible to grow unusually shaped single crystals.

Moving Cold Atoms with Quantum Ratchets

The nanoscale dimensions of biological motors make them susceptible to thermal noise, but such motors can produce force in one direction by alternating application of an asymmetric potential, or ratchet, with periods of thermal drift in the motion. The quantum version of such motors can operate without dissipation, as long as there is some means to break time-reversal symmetry in the system. **Salger *et al.*** (p. 1241) report on a coherent quantum ratchet device consisting of Bose-Einstein condensate cold atoms placed into an asymmetric sawtooth-potential created by optical lattices. Symmetry breaking was accomplished by phase shifts in the driving potentials. As expected for such a quantum ratchet, the current depended on the initial phase of the driving potential.

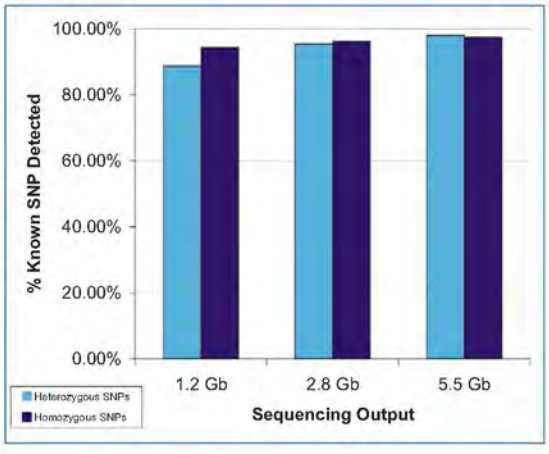
Cooperative Reduction

Selective redox transformation remains a general challenge in chemical synthesis. All too often, the most readily available precursor to a compound must be over-reduced (or over-oxidized) and then carefully coaxed back to a desired intermediate state. Such is the case with the synthesis of cyclohexanone, which is mass-produced for use in the preparation of nylon: Access by direct reduction of phenol is plagued by the rapid addition of too many hydrogen atoms to the substrate, producing an alcohol (cyclohexanol) in place of the ketone. **Liu *et al.*** (p. 1250) have discovered that the unexpected cooperation of supported palladium

Continued on page 1161

CREDITS (TOP TO BOTTOM): DAVID SHECHNER; LI *ET AL.*

NimbleGen SeqCap EZ Exome



SeqCap EZ Exome provides better coverage with reduced sequencing. Coding SNPs can be readily detected by combining SeqCap EZ Exome and next-gen sequencing at only a fraction of the cost of whole-genome sequencing.

Exome sequencing is an innovative technique to unravel the molecular basis for genetic disorders. By performing targeted sequencing on the most understood and functionally relevant 1% of the human genome, the exome, researchers can quickly and cost-effectively identify causative mutations in exons.

Roche NimbleGen is pleased to introduce the next-generation exome sequence capture technology, SeqCap EZ Exome.

- **Optimized performance for enhanced data with reduced sequencing.** 2.1 million optimized DNA probes (>10 probes/exon on average), provide greater uniformity resulting in substantial savings on sequencing compared with other methods.
- **Single-tube capture for an easy and scalable solution to enrich the exome.** New in-solution sequence capture method simplifies workflow, while optimized, platform-specific sequencing protocols require less time and starting material.
- **Built-in controls to evaluate capture performance.** Measure enrichment success with qPCR before sequencing.

Simplify your sequencing study today:
www.nimblegen.com/EZ

For life science research only.
Not for use in diagnostic procedures.

Roche NimbleGen, Inc.
Madison, WI USA

Roche

Continued from page 1159

and a Lewis acid such as aluminum trichloride—two catalysts widely used alone but rarely in concert—facilitates highly selective conversion of phenol to cyclohexanone near room temperature. The key appears to be inhibition of the undesired ketone-to-alcohol reduction step by the Lewis acid.

Untangling the Web

Chlorophyll-containing phytoplankton is at the core of the marine food web. **Martinez *et al.*** (p. 1253) combined satellite data about upper ocean chlorophyll and sea surface temperatures to demonstrate a clear connection between phytoplankton and sea surface temperatures on a multidecadal time scale. Basin-scale ocean dynamic processes such as the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation connect the physical, climate-related variability to changes in phytoplankton distribution and amount. Thus, improving the reliability of forecasts of large-scale ocean dynamics may help to improve predictions of changes in ocean community ecology.

Patterns of Change

The global climate record of the past 1500 years shows two long intervals of anomalous temperatures before the obvious anthropogenic warming of the 20th century: the warm Medieval Climate Anomaly between roughly 950 and 1250 A.D. and the Little Ice Age between around 1400 and 1700 A.D. It has become increasingly clear in recent years, however, that climate changes inevitably involve a complex pattern of regional changes, whose inhomogeneities contain valuable insights into the mechanisms that cause them. **Mann *et al.*** (p. 1256) analyzed proxy records of climate since 500 A.D. and compared their global patterns with model reconstructions. The results identify the large-scale processes—like El Niño and the North Atlantic Oscillation—that can account for the observations and suggest that dynamic responses to variable radiative forcing were their primary causes.

RSV in 3D

Respiratory syncytial virus (RSV) causes pneumonia and bronchiolitis in infants. RSV is an RNA virus in which the genomic RNA forms part of a nuclease-resistant helical ribonucleoprotein complex. **Tawar *et al.*** (p. 1279) now use x-ray and electron microscopy data to model the structure of this nucleocapsid complex and show how it can template RNA synthesis. The crystal structure shows RNA wrapped around a decameric ring of nucleocapsid protein. Combining this structure with electron microscopy data gives a model that shows how polymerase might read out the RNA bases without disassembling the nucleocapsid helix.

Inserting Introns

Introns—noncoding regions that interrupt coding gene sequences—are widespread throughout eukaryotic genomes, but intron gains and losses within and among species have been assumed to be rare. However, **W. Li *et al.*** (p. 1260) suggest that intron insertions can be relatively frequent within a population or species. By examining intron polymorphisms within genomes of different accessions of *Daphnia pulpex* (the water flea), and comparisons within the genus *Daphnia*, several instances of recent intron gains were observed, which appear to have occurred multiple times at the same site. Because intron insertions tend to be flanked by repetitive sequences, they may be the result of DNA damage repair mechanisms.

Pre-MicroRNA Export Machinery

Micro (mi) RNAs play a role in the regulation of many biological processes. Long transcripts are initially processed in the nucleus to yield pre-miRNAs that are translocated through the nuclear pore complex and further processed to mature miRNAs in the cytoplasm. **Okada *et al.*** (p. 1275; see the Perspective by **Stewart**) describe the crystal structure of pre-miRNA complexed with the exportin Exp5 and the small nuclear GTPase RanGTP. The structure shows that Exp5 and RanGTP protect the miRNA from degradation by nucleases, as well as facilitate transport to the cytoplasm. RNA recognition is mainly through ionic interactions that are sequence independent, and model-building suggests that this nuclear export machinery could accommodate other small-structured RNAs.

CREDIT: TAWAR *ET AL.*

CAS REGISTRY℠

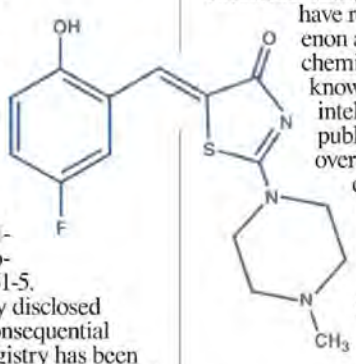
The "Gold Standard" of Substance Collections.

FROM THE EDITOR

A Scientific Milestone

This guest editorial is by **Matthew Toussant**, *senior vice president of editorial operations for CAS, who is responsible for the editorial production of CAS databases, principally CA and the CAS Registry.*

ON SEPT. 7, CAS scientists recorded the 50-millionth chemical substance into the CAS Registry. It received a unique identifier, the CAS Registry Number (CAS RN), and was associated with its authoritative source, in this case a World Intellectual Property Organization application, WO2009/097695, published on Aug. 13, 2009. The substance comes from the examples section of a 199-page patent document and is (5*Z*)-5-[(5-fluoro-2-hydroxyphenyl) methylene]-2-(4-methyl-1-piperazinyl)- 4(5*H*)-thiazolone, CAS RN 1181081-51-5.



The 50 million publicly disclosed substances represent a consequential milestone. The CAS Registry has been continuously operated for the purposes of uniquely identifying chemical substances since its inception, now more than 40 years ago. Surprisingly, it took CAS only nine months to register the last 10 million substances. In those nine months, CAS has registered at least 25 unique substances per minute.

It took 33 years for CAS to encounter the first 10 million substances in the published literature. As CAS forecast two years ago, the pace of discovery in chemistry, especially of small molecules, is increasing. And so, although in 2008 CAS registered a then-record total of 8.5 million substances, that record has already been shattered.

What other trends are evident on closer inspection of the Registry? Although CAS scientists review all forms of publications, we see that more than 60% of the new substances entering the CAS Registry are sourced from patents issued by a major patent office. Another significant percentage comes from commercial chemical catalogs. For some years, we have referred to this phenomenon as "the monetization of chemistry" because chemical knowledge embedded in the intellectual property of a public patent document has overtaken chemical knowledge in the form of more traditional published and shared media, such as scientific articles. How long will this trend of increasing protection of chemical knowledge persist? No one can predict; but for now the work of analyzing complex patents continues to challenge and bring out the best in the information analysis techniques and scientific expertise of CAS scientists and information technologists.

The CAS Registry is more than a simple list or compendium. It is to the chemist a lively and vast mosaic of chemical information that provides not only chemical names and vital literature references but ancillary information such as experimental and predicted property data, commercial availability, preparation details, spectra, and regulatory information from international sources.

Along with other key CAS databases, the CAS Registry can be an almost endless source of insight and a prod to creativity. The Registry has been called "the gold standard" of substance collections, and all of us at CAS, and our close partners around the world, are proud to have earned that recognition.

The CAS Registry is an embodiment of our mission as a division of ACS. That mission is to provide chemists and allied scientists with the world's best digital environment to search, retrieve, analyze, and link chemical information. At CAS, we aim to be comprehensive, and yet we apply careful standards before adding substances to the Registry. We seek to be timely, and yet we will not yield on quality standards and will commit to corrections when errors are found.

The CAS Registry reaches back more than 100 years and is adding new information at a prodigious rate. It is both the life's work of generations of CAS chemists and the beneficiary of the efforts of thousands of abstractors and indexers from dozens of countries. Many contributed vital work to CAS's information resources before the digital age and the conception of the Registry in the 1960s. CAS databases are an astonishing resource, and as we mark the milestone of 50 million substances, we look back in appreciation at this body of work from which we and world science benefit today.

Matthew Toussant

CAS is proud to reprint the *Chemical & Engineering News* guest editorial, marking an important milestone in the growth of research and scientific information.

To learn more about accessing CAS REGISTRY through our award-winning research tools, SciFinder® and STN®, visit www.cas.org.

Contact CAS at:
Phone: 800-753-4227 (North America)
614-447-3700 (worldwide)
E-mail: help@cas.org

www.cas.org

A division of the American Chemical Society

Bruce Alberts is Editor-in-Chief of *Science*.

On Incentives for Innovation

FOR SCIENCE TO THRIVE, IT IS CRUCIAL THAT THE SCIENTIFIC COMMUNITY ENCOURAGE THE BOLD ambitions and innovative spirit of young researchers. In my own area of science, the United States could do much more to support this important goal. U.S. biomedical science is a large and important research enterprise that currently includes over 100,000 graduate students and postdoctoral fellows. Of these, only a select few will go on to become independent research scientists in academia. Assuming that the system supporting this career path works well, these will be the individuals with the most talent and interest in such an endeavor: young people well positioned to make the scientific breakthroughs that societies need to survive and thrive. But the current system squanders the creativity and energy of these exceptionally gifted young people through a funding process that forces them to avoid risk-taking and innovation.

The traditional peer review system on which scientists depend for federal grant support values biomedical research projects that are almost certain to "work," encouraging young scientists to pursue a narrow range of projects that closely follow the proven paths of their mentors. As a result, many scientists pursue identical research ideas, creating a competition to finish and publish that can value speed over quality. Worse, the innovation that is essential for keeping science exciting and productive is replaced by a great deal of "me-too" science: research that has little chance of producing the breakthroughs needed to improve human health.

The U.S. National Institutes of Health (NIH) is by far the largest funder of research in this area, with a budget of about $30 billion per year. There, a scientist's proposal for research support is reviewed in about 200 "study sections"—with titles such as Anterior Eye Disease or Vaccines Against Microbial Diseases—each composed of scientists with a particular expertise. In 2007, frustration with how this generally conservative, risk-averse review process has dampened creative career paths led to the creation of the NIH Director's New Innovator Award. As one of many reviewers for this new program, I was asked to rank 35 10-page applications. To be eligible, an applicant must have received a doctoral degree no more than 10 years previously. Each investigator was asked to propose "highly innovative approaches to a high-impact problem," with no preliminary data required. In addition, an explanation was requested for why this work was unlikely to be funded through normal review mechanisms.

Often presenting bold, original ideas, these applications were a pleasure to read, and I thought that at least 4 of my 35 should have been funded. But with 2200 applications received for the 30 awards offered, not a single scientist on my list was selected. This year, a total of 54 of these 5-year grants were awarded, with projects that range from "A Biochip for Point-of-Care HIV/AIDS Diagnosis in the Developing World" to the "Intracellular Delivery and Targeting of Nanoparticles."*

The New Innovator Award and two others that specifically encourage innovation (NIH's Pioneer and Transformative R01 Awards) make a big difference to those who receive them. But there are far too few to change the culture for scientists starting new labs. Most remain unwilling to take the risk of pursuing ambitious ideas, recognizing that extensive preliminary results will be required to obtain funding from a traditional study section.

NIH has committed $80 million to support New Innovator Awards for 2010. One can ask whether this investment of only 0.27% of the NIH budget is appropriate. To me, the answer is a definite no. A major increase in the number of these 5-year awards to, say, 500 each year would send a very different signal to young people by supporting a culture of innovation and thus the long-term health of the scientific enterprise. Private foundations (for example, the Howard Hughes Medical Institute and Wellcome Trust) continue to promote initiatives that support creative and transformative research.** National governments should take serious note. **– Bruce Alberts**

10.1126/science.1184848

*http://nihroadmap.nih.gov/newinnovator/recipients09.asp. **J. Kaiser, *Science* **326**, 921 (2009)

EDITED BY GILBERT CHIN AND JAKE YESTON

ASTRONOMY

Unproductive Mergers

How do you make a massive galaxy that does not form new stars anymore? According to van der Wel *et al.* the answer is by merging galaxies of similar size. They selected 17,480 galaxies with low star-formation rates from the millions of galaxies mapped in the Sloan Digital Sky Survey and analyzed their shape distribution. Whereas the projected axial ratios of the less massive galaxies were consistent with a mix of shapes, from bulge- to disk-dominated galaxies, the most massive galaxies in their sample all turned out to be spheroidal. Because the only known mechanism capable of forming round galaxies is the merging of galaxies with mass ratios up to about 3, the authors concluded that such merging is the dominant mechanism for the production of massive, non–star-forming galaxies. — MJC

Astrophys. J. **706**, L120 (2009).

CHEMISTRY

Fluorine Diverted

The advent of frequency tunable laser sources briefly enticed chemists with the prospect of performing a sort of molecular surgery, whereby energy input at the precise resonance frequency of a bond could be used to break that bond, while leaving the rest of the molecule undisturbed. It quickly became clear, though, that most molecules wouldn't cooperate; the energy tended to spread across the whole molecular framework, despite being introduced at a specific site. Still, selective vibrational excitation would at worst have no impact. Surely it wouldn't hinder subsequent reactivity, would it? Very recently, precisely such a baffling outcome was observed in the reaction of fluorine atoms with trideuterated methane (CHD_3): Exciting the C-H bond actually reduced the likelihood of fluorine's breaking it, instead leading to an increased proportion of D-F product (see Zhang *et al.*, Reports, 17 July 2009, p. 303). Czakó and Bowman have now analyzed this reaction system theoretically. Preliminary examination of the potential energy surface suggested that vibration should lower the barrier toward breaking the excited bond, in keeping with intuition. Yet in quasiclassical trajectory simulations of low-energy collisions, inducing C-H bond vibration appeared instead to steer the F atom toward the deuterium centers, thus bolstering the experimental findings. — JSY

J. Am. Chem. Soc. **131**, 10.1021/ja906886z (2009).

ECOLOGY

Intrinsic Rhythms

Many plant species reproduce irregularly, flowering synchronously only every few years. Famous examples are the giant dipterocarp trees of Southeast Asian rainforests and the beech trees of temperate forests. Despite much research on the ecological and evolutionary aspects of this phenomenon, little is known about the proximate mechanisms by which the individual plants in a population achieve reliable synchrony in their flowering. Crone *et al.* addressed this question in an experimental study with a North American plant species, *Astragalus scaphoides*, which flowers in alternate years. By manipulating the resource budgets of individual plants (by reducing fruit set through flower removal), they showed that the levels of mobile carbohydrate control the frequency of flowering: Plants prevented from setting fruit maintained higher levels of resources and hence flowered again the next year; this resulted in pollen-limited reproduction, reducing seed set and leading to the reestablishment of synchrony the following year. — AMS

Ecol. Lett. **12**, 1119 (2009).

CELL BIOLOGY

The Scenic Route

The plasma membrane contains a large number of proteins and lipids, many of which are organized spatially into regions that mediate processes such as motility, adhesion, and signal transduction. Cilia are hairlike extensions that protrude from the plasma membrane and are found on the surface of most cells. They mediate mechanical and sensory interactions between the cell and its environment, as well as act as signaling centers for the Hedgehog and Wingless developmental pathways. In order to support these functions, proteins must be targeted to the ciliary membrane. For the transmembrane receptor pro-

CREDITS (TOP TO BOTTOM): NASA; ELIZABETH.CRONE@CFC.UMT.EDU

tein Smoothened, which is part of the Hedgehog pathway and localizes to the cilium upon activation, current models have focused on its delivery by membrane vesicles moving from the Golgi apparatus to the cell surface. Milenkovic *et al.* have identified an alternative route: lateral transport within the plasma membrane. They found that Smoothened could move to the ciliary membrane from the bulk plasma membrane when the Hedgehog pathway was activated. Increasing the levels of the second-messenger molecule cAMP, which is involved in lateral transport in algae, could also induce lateral transport of Smoothened, suggesting a conserved mechanism of regulation. Perhaps to improve efficiency and accuracy, Smoothened is transported to the ciliary membrane along two different routes. — HP

J. Cell Biol. **187**, 365 (2009).

MICROBIOLOGY

Journeys of a Pathogen

Leprosy is caused by *Mycobacterium leprae*, which infects the neurons and mucosa of humans. Strikingly, its genome has decayed extensively, which probably accounts for its slow rate of growth; global variation between strains is limited to four SNP clusters, each having a distinctive geographical prevalence that reflects past human migrations. Monot *et al.* have sequenced *M. leprae* strains from Brazil, Thailand, and the United States, and made comparisons with archaeological samples. Genotyping the pathogen can be used to pinpoint a person's origins quite precisely. For example, samples from an Egyptian burial revealed that the corpse had been infected with a European strain of leprosy and was more probably a Roman legionary than an African. Their updated map of leprosy strains supports the out-of-Africa theory for modern humans and indicates that leprosy crossed to Asia twice: once via a southern route into India and through Indonesia to the Philippines, and also via a northern route along the Silk Road in countercurrent fashion to the Black Death. Subsequently, leprosy arrived in South America via European immigrants rather than by pre-Columbian colonization via the Bering Strait. — CA

Nat. Genet. **41**, 10.1038/ng.477 (2009).

GEOLOGY

Fresh from the Ocean

Freshwater springs have been found along the Atlantic coastal shelves of North America and Europe even at depths of several hundred meters below sea level, and several hundred kilometers offshore. In some areas (for example, the southern United States), springs are emanating from coherent aquifers that are recharged high enough above sea level to provide the gravitational potential to push the fresh water out into the denser saltwater. Elsewhere though, the fresh water seems to have been trapped during glacial cycles when sea level was considerably lower. Cohen *et al.* provide an overview of these settings and explore their formation and resource potential through a series of numerical models. Their models suggest that extensive recharging occurred beneath Pleistocene ice sheets, which extended out onto the shelves or along submarine canyons. They estimate that some submarine aquifers off New England may contain as much as 10,000 km^3 of fresh water. Globally, there may be as much as 300,000 km^3 of fresh water, though it is nonrenewable. — BH

Groundwater 10.1111/j.1745-6584.2009.00627.x (2009).

PSYCHOLOGY

Short-Circuiting the Path to Violence

Self-esteem and narcissism are sometimes invoked as explanations for outbreaks of classroom aggression, whether physical or verbal. That is, too little of one or too much of the other putatively contribute to the tendency to lash out when threatened, and hence there have been efforts to boost students' self-esteem as a means of reducing the frequency of confrontations. Thomaes *et al.* argue instead that the critical need is to buffer self-esteem—whether a student has a lot or a little of it—and they conducted a randomized field trial in the Netherlands to assess the efficacy of a brief affirmation exercise designed to buttress self-esteem. Before the intervention, aggressive behavior was prevalent among the individuals who scored high on the narcissism scale and low on self-esteem. Afterward, acts of aggression decreased even though the individual levels of self-esteem (both high and low) were unchanged, supporting the conclusion that firming up one's sense of worth reduces the vulnerability to threat and thus diminishes the motivation to launch defensive counterattacks. — GJC

Psychol. Sci. **20**, 10.1111/j.1467-9280.2009.02478.x (2009).

CREDIT: ISTOCKPHOTO.COM

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6581, FAX 202-371-9227
Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES please call 202-326-6755 for any questions or information

REPRINTS: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578

PERMISSIONS 202-326-7074, FAX 202-682-0816

MEMBER BENEFITS AAAS/Barnes&Noble.com bookstore www.aaas.org/bn; AAAS Online Store www.apisource.com/aaas/ code MKB6; AAAS Travels: Betchart Expeditions 800-252-4910; Apple Store www.apple.com/eppstore/aaas; Bank of America MasterCard 1-800-833-6262 priority code FAA3YU; Cold Spring Harbor Laboratory Press Publications www.cshlpress.com/affiliates/aaas.htm; GEICO Auto Insurance www.geico.com/landingpage/go51.htm?logo=17624; Hertz 800-654-2200 CDP#343457; Office Depot https://bsd.officedepot.com/portalLogin.do; Seabury & Smith Life Insurance 800-424-9883; Subaru VIP Program 202-326-6417; VIP Moving Services www.vipmayflower.com/domestic/index.html; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

science_editors@aaas.org	(for general editorial queries)
science_letters@aaas.org	(for queries about letters)
science_reviews@aaas.org	(for returning manuscript reviews)
science_bookrevs@aaas.org	(for book review queries)

Published by the American Association for the Advancement of Science (AAAS), *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science, engineering, and innovation throughout the world for the benefit of all people. The goals of the association are to: enhance communication among scientists, engineers, and the public; promote and defend the integrity of science and its use; strengthen support for the science and technology enterprise; provide a voice for science on societal issues; promote the responsible use of science in public policy; strengthen and diversify the science and technology workforce; foster education in science and technology for everyone; increase public engagement with science and technology; and advance international cooperation in science.

INFORMATION FOR AUTHORS
See pages 807 and 808 of the 6 February 2009 issue or access www.sciencemag.org/about/authors

EDITOR-IN-CHIEF **Bruce Alberts**
EXECUTIVE EDITOR **Monica M. Bradford**
NEWS EDITOR **Colin Norman**
MANAGING EDITOR, RESEARCH JOURNALS **Katrina L. Kelner**
DEPUTY EDITORS **R. Brooks Hanson, Barbara R. Jasny, Andrew M. Sugden**

EDITORIAL SENIOR EDITORS/COMMENTARY Lisa D. Chong, Brad Wible; SENIOR EDITORS Gilbert J. Chin, Pamela J. Hines, Paula A. Kiberstis (Boston), Marc S. Lavine (Toronto), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Phillip D. Szuromi (Tennessee), Valda Vinson, Jake S. Yeston; ASSOCIATE EDITORS Kristen L. Mueller, Jelena Stajic, Nicholas S. Wigginton, Laura M. Zahn; RESEARCH ASSOCIATE Alexis Wynne Mogul; ONLINE EDITOR Stewart Wills; ASSOCIATE ONLINE EDITORS Robert Frederick, Tara S. Marathe; WEB CONTENT DEVELOPER Martyn Green; BOOK REVIEW EDITOR Sherman J. Suter; ASSOCIATE LETTERS EDITOR Jennifer Sills; EDITORIAL MANAGER Cara Tate; SENIOR COPY EDITORS Jeffrey E. Cook, Cynthia Howe, Harry Jach, Barbara P. Ordway, Trista Wagoner; COPY EDITORS Chris Filiatreau, Lauren Kmec; EDITORIAL COORDINATORS Carolyn Kyle, Beverly Shields; PUBLICATIONS ASSISTANTS Ramatoulaye Diop, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Scott Miller, Jerry Richardson, Jennifer A. Seibert, Brian White, Anita Wynn; EDITORIAL ASSISTANTS Emily Guise, Michael Hicks, Patricia M. Moore; EXECUTIVE ASSISTANT Sylvia S. Kihara; ADMINISTRATIVE SUPPORT Maryrose Madrid; EDITORIAL FELLOW Melissa R. McCartney

NEWS DEPUTY NEWS EDITORS Robert Coontz, Eliot Marshall, Jeffrey Mervis, Leslie Roberts; CONTRIBUTING EDITORS Elizabeth Culotta, Polly Shulman; NEWS WRITERS Yudhijit Bhattacharjee, Adrian Cho, Jennifer Couzin, David Grimm, Constance Holden, Jocelyn Kaiser, Sam Kean, Richard A. Kerr, Eli Kintisch, Andrew Lawler (New England), Greg Miller, Elizabeth Pennisi, Robert F. Service (Pacific NW), Erik Stokstad; INTERNS Michael Torrice, Jue Wang; CONTRIBUTING CORRESPONDENTS Dan Charles, Jon Cohen (San Diego, CA), Daniel Ferber, Ann Gibbons, Robert Koenig, Mitch Leslie, Charles C. Mann, Virginia Morell, Gary Taubes; COPY EDITORS Linda B. Felaco, Melvin Gatling, Melissa Raimondi; ADMINISTRATIVE SUPPORT Scherraine Mack, Fannie Groom; BUREAUS New England: 207-549-7755; San Diego, CA: 760-942-3252, FAX 760-942-4979; Pacific Northwest: 503-963-1940

PRODUCTION DIRECTOR James Landry; SENIOR MANAGER Wendy K. Shank; ASSISTANT MANAGER Rebecca Doshi; SENIOR SPECIALISTS Steve Forrester, Chris Redwood; SPECIALIST Anthony Rosen; PREFLIGHT DIRECTOR David M. Tompkins; MANAGER Marcus Spiegler; SPECIALIST Jason Hillman

ART DIRECTOR Yael Kats; ASSOCIATE ART DIRECTOR Laura Creveling; SENIOR ILLUSTRATORS Chris Bickel, Katharine Sutliff; ILLUSTRATOR Yana Greenman; SENIOR ART ASSOCIATES Holly Bishop, Preston Huey, Nayomi Kevitiyagala; ART ASSOCIATES Jessica Newfield, Matthew Twombly; PHOTO EDITOR Leslie Blizard

SCIENCE INTERNATIONAL

EUROPE (science@science-int.co.uk) EDITORIAL: INTERNATIONAL MANAGING EDITOR Andrew M. Sugden; SENIOR EDITOR/COMMENTARY Julia Fahrenkamp-Uppenbrink; SENIOR EDITORS Caroline Ash, Stella M. Hurtley, Ian S. Osborne, Peter Stern; ASSOCIATE EDITOR Maria Cruz; LOCUM EDITOR Helen Pickersgill; EDITORIAL SUPPORT Deborah Dennison, Rachel Roberts, Alice Whaley; ADMINISTRATIVE SUPPORT John Cannell, Janet Clements, Louise Moore; NEWS: EUROPE NEWS EDITOR John Travis; DEPUTY NEWS EDITOR Daniel Clery; CONTRIBUTING CORRESPONDENTS Michael Balter (Paris), John Bohannon (Vienna), Martin Enserink (Amsterdam and Paris), Gretchen Vogel (Berlin)

ASIA Japan Office: Asca Corporation, Tomoko Furusawa, Rustic Bldg. 7F, 77 Tenjin-cho, Shinjuku-ku, Tokyo 162-0808, Japan; +81 3 6802 4616, FAX +81 3 6802 4615, inquiry@sciencemag.jp; ASIA NEWS EDITOR Richard Stone (Beijing: rstone@aaas.org); CONTRIBUTING CORRESPONDENTS Dennis Normile [Japan: +81 (0) 3 3391 0630, FAX +81 (0) 3 5936 3531; dnormile@gol.com]; Hao Xin [China: +86 (0) 10 6307 4439 or 6307 3676, FAX +86 (0) 10 6307 4358; cindyhao@gmail.com]; Pallava Bagla [South Asia: +91 (0) 11 2271 2896; pbagla@vsnl.com]

EXECUTIVE PUBLISHER **Alan I. Leshner**
PUBLISHER **Beth Rosner**

FULFILLMENT SYSTEMS AND OPERATIONS (membership@aaas.org); DIRECTOR Waylon Butler; SENIOR SYSTEMS ANALYST Nomuna Nyamaa; CUSTOMER SERVICE SUPERVISOR Pat Butler; SPECIALISTS Latoya Casteel, LaVonda Crawford, Vicki Linton, April Marshall; DATA ENTRY SUPERVISOR Cynthia Johnson; SPECIALISTS Shirlene Hall, Tarrika Hill, William Jones

BUSINESS OPERATIONS AND ADMINISTRATION DIRECTOR Deborah Rivera-Wienhold; ASSISTANT DIRECTOR, BUSINESS OPERATIONS Randy Yi; MANAGER, BUSINESS ANALYSIS Eric Knott; MANAGER, BUSINESS OPERATIONS Jessica Tierney; FINANCIAL ANALYSTS Priti Pamnani, Celeste Troxler; RIGHTS AND PERMISSIONS: ADMINISTRATOR Emilie David; ASSOCIATE Elizabeth Sandler; MARKETING DIRECTOR Ian King; MARKETING MANAGERS Allison Pritchard, Alison Chandler, Julianne Wielga; MARKETING ASSOCIATES Aimee Aponte, Mary Ellen Crowley, Adrian Parham, Wendy Wise; MARKETING EXECUTIVE Jennifer Reeves; DIRECTOR, SITE LICENSING Tom Ryan; DIRECTOR, CORPORATE RELATIONS Eileen Bernadette Moran; PUBLISHER RELATIONS, ERESOURCES SPECIALIST Kiki Forsythe; SENIOR PUBLISHER RELATIONS SPECIALIST Catherine Holland; PUBLISHER RELATIONS, EAST COAST Phillip Smith; PUBLISHER RELATIONS, WEST COAST Philip Tsolakidis; FULFILLMENT SUPERVISOR Iquo Edim; FULFILLMENT COORDINATOR Carrie MacDonald; MARKETING MANAGER Christina Schlecht; MARKETING ASSOCIATE Mary Lagnaoui; ELECTRONIC MEDIA: MANAGER Lizabeth Harman; PROJECT MANAGER Trista Snyder; ASSISTANT MANAGER Lisa Stanford; SENIOR PRODUCTION SPECIALISTS Ryan Atkins, Christopher Coleman, Walter Jones; PRODUCTION SPECIALISTS Nichele Johnston, Kimberly Oster

ADVERTISING DIRECTOR, WORLDWIDE AD SALES Bill Moran

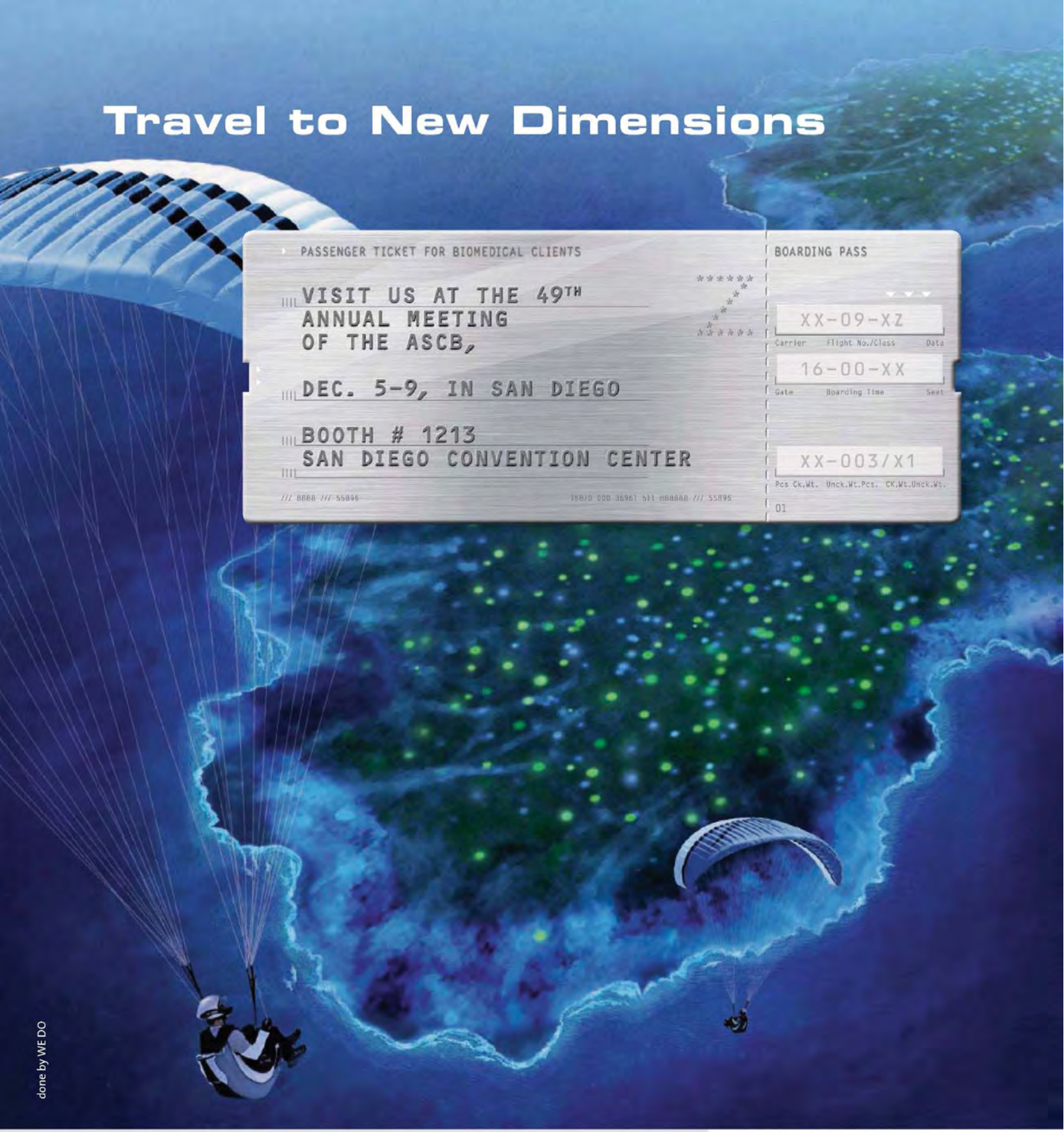
PRODUCT (science_advertising@aaas.org); MIDWEST/WEST COAST/W. CANADA Rick Bongiovanni: 330-405-7080, FAX 330-405-7081; EAST COAST/E. CANADA Laurie Faraday: 508-747-9395, FAX 617-507-8189; UK/EUROPE/ASIA Roger Goncalves: TEL/FAX +41 43 243 1358; JAPAN ASCA Corporation, Nanako Ide +81 (0) 3 6802 4616, FAX +81 (0) 3 6802 4615; ads@sciencemag.jp; SENIOR TRAFFIC ASSOCIATE Deiandra Simms

COMMERCIAL EDITOR Sean Sanders: 202-326-6430

PROJECT DIRECTOR, OUTREACH Brianna Blaser

CLASSIFIED (advertise@sciencecareers.org); U.S.: SALES MANAGER Daryl Anderson: 202-326-6543; MIDWEST Tina Burks: 202-326-6577; EAST COAST Alexis Fleming: 202-326-6578; WEST/SOUTH CENTRAL Nicholas Hintibidze: 202-326-6533; SALES COORDINATORS Rohan Edmonson, Shirley Young; INTERNATIONAL: SALES MANAGER Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532; SALES Susanne Kharraz, Dan Pennington, Alex Palmer; SALES ASSISTANT Lisa Patterson; JAPAN ASCA Corporation, Jie Chin +81 (0) 3 6802 4616, FAX +81 (0) 3 6802 4615; careerads@sciencemag.jp; ADVERTISING SUPPORT MANAGER Karen Foote: 202-326-6740; ADVERTISING PRODUCTION OPERATIONS MANAGER Deborah Tompkins; SENIOR PRODUCTION SPECIALIST/GRAPHIC DESIGNER Amy Hardcastle; SENIOR PRODUCTION SPECIALIST Robert Buck; SENIOR TRAFFIC ASSOCIATE Christine Hall

Travel to New Dimensions



done by WE DO

Book your ticket to new insights today. It has never been as easy to make discoveries beyond the boundaries of current knowledge. Carl Zeiss works together with researchers around the world to develop systems that open up new perspectives in the life sciences. Find out more about ELYRA, LSM 780 and VivaTome – three unique new vehicles to take you to new dimensions.

www.zeiss.de/nd_science

ZEISS

We make it visible.

Your Power for Health

greiner bio-one

Give your Samples an Identity

Cryo.s™ with Datamatrix Code from Greiner Bio-One

- Fast, easy and secure processing of sample information
- Robust, laser-written symbols ensure data safety even after long-term storage
- 100% readability for each Datamatrix
- Encodes up to 24 numeric and 16 alphanumeric characters
- Well-established tube quality with more than 20 years of experience in cryo storage

Germany (Main office): Greiner Bio-One GmbH, info@de.gbo.com I **Austria:** Greiner Bio-One GmbH, office@at.gbo.com
Belgium: Greiner Bio-One BVBA/SPRL, info@be.gbo.com I **Brazil:** Greiner Bio-One Brasil, office@br.gbo.com I **China:** Greiner Bio-One GmbH, info@cn.gbo.com
France: Greiner Bio-One SAS, infos@fr.gbo.com I **Japan:** Greiner Bio-One Co. Ltd., info@jp.gbo.com I **Netherlands:** Greiner Bio-One B.V., info@nl.gbo.com
UK: Greiner Bio-One Ltd., info@uk.gbo.com I **USA:** Greiner Bio-One North America Inc., info@us.gbo.com

www.gbo.com/bioscience

How's Your Gaydar?

Can you guess a woman's sexual orientation just from her face? Surprisingly, your guess would be better than flipping a coin.

Psychologist Nicholas Rule of Tufts University in Medford, Massachusetts, and colleagues asked 21 college students the same question about 192 photos—cropped to eliminate hair and ears—of gay and straight women from dating Web sites. The undergraduates guessed right 64% of the time and scored better than chance—53%—even when they saw only the women's eyes, the researchers report this month in the *Journal of Experimental Social Psychology*. In 2008, Rule reported similar results with male faces.

The process appears to be unconscious, Rule says. Subjects were more accurate when told to make snap judgments than when they pondered their decisions. Rule believes subtle differences in facial muscles caused by habitual expressions may be the clue. Previous work has shown that homosexuals tend to adopt facial expressions more typical of the opposite sex. The results show that we know much more about others from snippets of information than we realize, says psychologist David Kenny of the University of Connecticut, Storrs.

In Search of Darkness

Really dark skies are becoming a rarity these days. Galloway Forest Park in Scotland was officially named as the United Kingdom's first "Dark Sky Park" last week at the meeting of the International Dark-Sky Association in Phoenix, Arizona.

Now there's a Web site to guide amateur astronomers. ClearDarkSky.com, created by software designer Attilla Danko of Ottawa, Canada, pinpoints the best nights for viewing the sky at more than 3800 locations in the United States, Canada, and Mexico.

The site translates forecast maps from the Canadian Meteorological Centre into diagrams showing conditions for the next 48 hours. They include hour-by-hour predictions of factors that affect viewing transparency: cloud cover, the amount of water vapor in the air, and seeing, an indicator of turbulence and atmospheric temperature differences.

The site also forecasts wind, temperature, and humidity at ground level, which influences whether eyepieces and lenses will collect condensation. If your locale isn't on the list, you can nominate it.

Sclerotic Egyptians

Pharaohs were worshiped as gods, but their hearts were all too human. Cardiologists and Egyptologists recently used CT scanners to scrutinize the mummies of 16 ancient Egyptians and found evidence of heart disease as far back as 1530 B.C.E.

Because mummification often involved removing the heart, the scientists examined mostly arteries and tracks where arteries used to run. They found both coated with calcium gunk, a sign of atherosclerosis. Scientists have dissected mummies to look for signs of diseases before, including leprosy, arthritis, and heart disease, but never with modern nondestructive imaging techniques.

The work, reported on 18 November in the *Journal of the American Medical Association*, shows that, given the right environment, people have always been susceptible to heart trouble, says co-author Samuel Wann, a cardiologist at the Wisconsin Heart Hospital in Milwaukee. In this case, the mummies came from the highest social castes of Egypt, which often dined on fatty delicacies. Nevertheless, given the less-sedentary lives of most ancient people, Wann says the findings are surprising: "I would have thought that atherosclerosis was a disease of modern man, not something that affected people in the time of Moses." Sandra Olsen, curator of anthropology at the Carnegie Museum of Natural History in Pittsburgh, Pennsylvania, says the work "throws out a lot of old myths about how we're going to pot because of our modern lifestyle."

MUSTINESS MARKER

The evocative smell of old books is the odor of decay. Analytical chemist Matija Strlič and colleagues at University College London have chemically analyzed the smell—which they describe as "a combination of grassy notes with a tang of acids and a hint of vanilla over an underlying mustiness"—and find that its composition is associated with paper degradation.

The researchers studied 72 samples of paper produced between 1850 and 1990, a period when paper was highly acidic. With liquid chromatography–mass spectrometry, they assessed the condition and stability of the paper through the volatile organic compounds it emitted. They also identified two compounds, lignin and rosin, that are associated with rapid degradation, they reported last month in *Analytical Chemistry*. Strlič hopes the work can eventually help conservators choose strategies for preserving books.

High-acid paper from 1909.

Barry Knight, head of Conservation Research at the British Library in London, says similar approaches might replace more invasive techniques for gauging the status of other historical items such as oil paintings and natural history specimens. Ultimately, Strlič envisions doing the job with small hand-held devices. Such a device, says Knight, "could sniff the book as it came up to the [library] reading room, ... and if it turned out that it was in bad condition, it could ring a bell" to alert a librarian.

CREDITS (TOP TO BOTTOM): AP PHOTOS; AP PHOTOS; MATIJA STRLIČ

HIV infections climb in Asia
1174

Q&A with Brazil's environment minister
1175

BIOTECHNOLOGY

Bankruptcy Won't Stop deCODE, Says Its Founder, Stefánsson

When the Icelandic company deCODE genetics Inc. filed for bankruptcy last week, many observers saw it as a bleak illustration of how hard it is to profit from research on the human genome. But deCODE and its flamboyant leader, Kari Stefánsson, aren't going away just yet. The company will continue operations with a loan from a potential buyer. And Stefánsson says he hopes to stay on as leader of deCODE's hunt for disease-causing genes.

Stefánsson founded deCODE in 1996 based on the controversial notion of collecting genetic data on Iceland's citizens, combining them with medical and genealogical data and mining this biobank for disease markers. Although deCODE's research has been a roaring success, generating many high-impact papers, the company's efforts to make money from new drugs and genetic tests have failed. For the past year, deCODE has been offering some of its components for sale.

What next? Kari Stefánsson is planning a new genetics firm called ... deCODE.

Last summer, deCODE was in discussions with the Wellcome Trust, a U.K. biomedical charity, about taking over support for the biobank (*Science*, 28 August, p. 1054). No deal emerged, however. Stefánsson denies rumors that deCODE's restriction on data sharing was an issue. The "fundamental reason" was that Wellcome wanted to run deCODE as a nonprofit research institution, which was not possible because of the company's legal obligations to pay a return to investors, he says. However, he "was moved by the way they reached out to us."

Instead, the company is trying a different way to survive. Last week, deCODE's U.S. parent company filed for Chapter 11 bankruptcy in Delaware, which means it will reorganize its debt and keep afloat for now. The company has a $14 million offer from Saga Investments for its drug-development and -discovery programs and for Islensk Erfdagreining, the deCODE subsidiary in Reykjavík that runs the biobank and genetic testing services. The parent company expects to be liquidated, however, and stockholders are unlikely to get any of their money back.

The deal must first be approved by a court, and other bidders could step in. Stefánsson says that if Saga's offer prevails—he hopes to know by early January—he expects to serve as executive chair of a new company and president of research. Someone else will serve as CEO and lead commercial operations. "It will be called deCODE," he says.

Stefánsson's research plan is to move on to the next stage in the hunt for disease genes: the search for rare variants. It is thought that these variants may confer higher disease risk than those found so far and would provide new insights into the biology of disease. Stefánsson says that by mid-2011 he plans to sequence the complete genomes of 2500 individuals, which, combined with genealogical data, should be enough to find variants occurring in 0.1% of Iceland's entire population of 320,000 people.

Some researchers say that because it is easier to find rare variants in a homogeneous population like Iceland's, the company stands a fighting chance of remaining a leader in this area. Disease-gene hunters have "tacked back to where the field started," says Stephen Chanock, a geneticist at the U.S. National Cancer Institute. **–JOCELYN KAISER**

FUSION

Schedule Concerns Delay ITER's Go-Ahead

The scientific and engineering team building the ITER fusion reactor failed to win an expected endorsement from the project's governing council last week. The council, which represents the seven international partners in the project—China, the European Union, India, Japan, South Korea, Russia, and the United States—sent the team back to do more work on the proposed construction schedule for the mammoth undertaking.

ITER is an experimental reactor that aims to show that nuclear fusion, the power source of the sun and stars, could be harnessed to generate energy on Earth. A site has been cleared at Cadarache in southern France for construction, and ITER staff have been racing for months to get the final project baseline documents, which describe the design, cost estimates, and planned schedule, ready for the 18–19 November council meeting at Cadarache (*Science*, 13 November, p. 932). But some council members voiced concern that the schedule, which aimed to start the reactor by 2018, was not realistic and that there was too high a risk that some part of the immensely complicated effort could go wrong.

A slip in the schedule would invariably mean increased costs, and the council is already concerned about budget estimates, which, sources say, may have doubled from €5 billion since the partners signed up in 2006. So the council told ITER staff to nail down more firmly the risks, both technical and organizational, involved in the schedule and come back in February with earliest and latest possible start-up dates. "Europe is very concerned about the risk of pushing ahead too fast," says Steven Cowley, head of the Culham Centre for Fusion Energy in Abingdon, U.K. "Building this is arguably harder than building the [Large Hadron Collider] because everything is inside everything else. You better get it right." Discussion of the cost estimates appears to have been put aside until the schedule has been resolved.

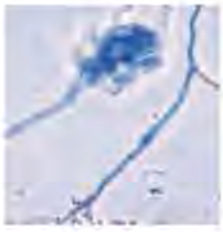
–DANIEL CLERY

CREDIT: IAN LANGSDON/REUTERS

INFECTIOUS DISEASES

Farm Fungicides Linked to Resistance in a Human Pathogen

A team of Dutch researchers has reignited a debate on the agricultural use of fungicides with a review in the December issue of *The Lancet Infectious Diseases*. The authors maintain that the massive use of fungicides to protect European orchards, vineyards, and grain fields may be contributing to resistance against drugs used to treat people with life-threatening infections of *Aspergillus fumigatus*. Although the overuse of antibiotics in animal husbandry is known to have caused resistance in the human population, this would be the first time a similar link is found between farm use of fungicides and human health.



Deadly mold. Invasive *Aspergillus fumigatus* infections can be fatal.

If true, the authors warn ominously, that "confronts us with a major challenge with worldwide dimensions." But Herbert Hof, director of the Institute for Medical Microbiology at the University of Heidelberg in Germany, accuses them of crying wolf, saying the paper amounts to "publicity seeking" by frightening the public "in the way horror films do." The group does have its supporters, however. David Denning of the University of Manchester, who heads the United Kingdom's National Aspergillosis Centre, says "they have a very strong case." The leader of the team, Paul Verweij of Radboud University Nijmegen Medical Centre in the Netherlands, concedes that they haven't yet clinched the case, but he says enough evidence has accumulated to issue a warning.

A. fumigatus causes infections, sometimes fatal, primarily in people with compromised immune systems and certain diseases, such as chronic obstructive pulmonary disease. Patients are thought to become infected when they inhale spores of the fungus, which are ubiquitous in soil. Drugs of a class called azoles are doctors' mainstay, and resistance has long been known to crop up in individual patients. The mutations in the fungus that cause the resistance usually differ from one patient to the next; in a paper published in July, for instance, Denning's team reported finding 18 different mutations in an *Aspergillus* gene called *cyp51A* in 30 patients in the United Kingdom.

But Verweij's team has found something strange in resistant *Aspergillus* strains in the Netherlands: In 94% of the isolates from his own hospital and 69% of those from other Dutch hospitals, the resistance was caused by a single pair of mutations—a point mutation in *cyp51A* and a so-called tandem repeat in the gene's promoter. To Verweij, that similarity points to a new scenario: that all the patients breathed in spores that were already resistant. That's why he believes there's an environmental cause.

Azoles are used to ward off a range of plant pathogens and are applied on 50% of Europe's grain and grape acreage, says plant pathologist Gert Kema of Wageningen University in the Netherlands, a co-author of the paper. Much smaller amounts are used in the United States, where farming is less-intensive and spraying is less cost-effective. But the compounds are popular in other parts of the world as well, he says.

The risk that heavy agricultural use of azoles might lead to resistance problems in people has been debated for years. In 2002, an expert panel for the European Commission concluded that it was unlikely. But since then, the evidence has been building, Verweij says. The resistant fungus found in patients is also resistant to certain agricultural fungicides, which is suggestive of a link. And in a paper published in June, his group showed that resistant *Aspergillus* could be isolated from soil in flower beds close to hospitals and in commercial compost, leaves, and seeds bought at a garden center. Thirteen of 15 of these environmental samples also had the two mutations seen in clinical isolates.

In an e-mail to *Science*, Hof called the authors "prejudiced" and said resistance in fungi is unlikely to become a major public health problem because unlike bacteria, fungi don't swap resistance genes. Dominique Sanglard of the University Hospital in Lausanne, Switzerland, on the other hand, says the Dutch researchers seem to be on to something real, although many questions remain. A key step is to show that one or more of the azole fungicides—at least 30 of them are on the market—can actually trigger the mutations in *A. fumigatus* seen in hospitals, says Verweij; that study is already under way.

And what if the link is proven? Verweij says a ban on certain fungicides could be an option. The team has been talking to several fungicide producers, and "they aren't very keen on studying this further," he says. A spokesperson for Syngenta, a major azole fungicide producer, says that resistance may have arisen in other ways—such as azole use in cosmetics—and that the company is "not convinced" of a causal link.

–MARTIN ENSERINK

Farmers' friend. Azoles are used to protect a wide variety of crops from fungi.

CREDITS (TOP TO BOTTOM): CDC; HENRY ARDEN/GETTY IMAGES

AIDS

Asia Grapples With Unexpected Wave of HIV Infections

When physician Edsel Salvaña discovered earlier this year that new AIDS cases and HIV infections at his clinic in the Philippine General Hospital in Manila were running nearly double the 2008 rate, he was concerned. When colleagues at other clinics began reporting the same phenomenon, he grew alarmed. "It seems like HIV is starting to get out of hand," Salvaña says.

Be safe. The Philippine Red Party kicked off its HIV awareness campaign with a benefit concert.

Not just in the Philippines. Across Asia-Pacific, a region so far largely spared the worst of the AIDS epidemic, HIV infections are rising rapidly, primarily among men having sex with men (MSM). In Hong Kong, Taiwan, Singapore, and Japan, where reporting is mandatory and trends can be tracked over time, HIV infections among MSM more than doubled between 2003 and 2007. Studies cited in "The global epidemic of HIV infection among men who have sex with men" in the July issue of *Current Opinion in HIV and AIDS* by Frits van Griensven, an epidemiologist at the U.S. Centers for Disease Control and Prevention's Bangkok office, and colleagues put HIV infection rates in MSM at about 30% in Bangkok and Yangon, more than 15% in parts of India, and more than 5% in Beijing. In the Philippines, the number of new HIV infections rose from 210 in 2005 to 549 in the first 9 months of 2009, with the proportion of those infected through homosexual contact jumping from less than 40% to more than 70%. The numbers "are pretty frightening," says van Griensven, who says the statistics capture only a portion of those actually infected. In the Philippines, says Salvaña, most infections are identified only after an individual develops AIDS symptoms.

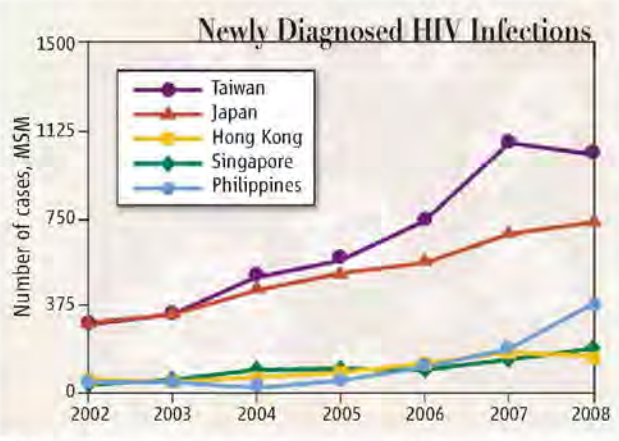
Recognizing a worsening problem, the World Health Organization (WHO) in late September decried the "slow, fragmented and insufficient" anti-HIV efforts in the Western Pacific. With this concern in mind, the U.S. Agency for International Development and the United Nations Development Programme this week sponsored a meeting in Bangkok to bolster the region's HIV prevention, treatment, and care for MSM.

The situation is likely to get worse before it gets better. Due to "significant levels of stigma, discrimination, and criminalization [of male homosexuality], there has been almost no investment" in some countries for interventions targeting this group, says Shivananda Khan, who heads the Naz Foundation International in Lucknow, India. "It is a perfect scenario for a major concentrated epidemic."

Unlike the AIDS epidemic in sub-Saharan Africa that has battered whole communities, Asia-Pacific cases are concentrated in three high-risk groups: IV injecting drug users, commercial sex workers, and MSM. "It took several years for the international community to fully understand why," says Massimo Ghidinelli, WHO Western Pacific regional adviser in HIV/AIDS and Sexually Transmitted Infections. Behavioral studies and computer modeling indicate that the epidemic in Africa is sustained by men and women who often have multiple concurrent partners and change partners frequently, Ghidinelli explains. Throughout much of Asia-Pacific, he says, individuals may have several partners over time, but they typically end one relationship before starting another. And in parts of Asia, women tend to have one or at most a few partners throughout their lives. "The dynamics to maintain and amplify an epidemic [outside the risk groups] are not there," Ghidinelli says. Another factor keeping infection rates low in the Philippines is widespread circumcision, which evidence shows inhibits infection.

Although national HIV infection rates are still very low in Asia-Pacific, in MSM they are skyrocketing. The rates are "similar to how quickly they went up among female sex workers or injecting drug users 20 years ago," says van Griensven. Even if the epidemic is confined to high-risk groups, Ghidinelli says, rising numbers of HIV-infected individuals could swamp health services and strain budgets of developing countries.

Cultural factors complicate the issue. Some countries, such as Thailand, tolerate homosexuality, whereas in Malaysia and elsewhere it is a criminal offense. Although prosecutions are rare, laws banning homosexual acts make it difficult to fund interventions. Another factor is that in several Asia-Pacific countries, men who dress and act like women are common partners for men who don't consider themselves homosexual, Khan says. Denial and social marginalization mean that only an estimated 4% of the region's governmental spending on HIV/AIDS targets MSM, and interventions reach only 9% of this population. A 2008 report by the United Nations Joint Programme on HIV/AIDS forecast that without more aggressive measures, MSM could become the primary reason for new HIV infections in Asia, reaching a million cases a year by 2020.



Taking off. Unprotected male-to-male sex is driving an alarming rise in HIV infections in Asia-Pacific.

The situation is not grim everywhere. Some countries are reporting a leveling off of new HIV infections, says van Griensven. Last year, he says, Hong Kong and Taiwan cut new infections through "a bombardment of interventions." Under Hong Kong's HIV/AIDS strategy adopted in 2007, the government expanded its own efforts and funded community organizations to study homosexual behavior, promote risk awareness and safer sex, and launch media campaigns to erase the stigma of being HIV-positive.

Although some governments are waking up to the problem, Khan says, others are in denial. The Philippines is coming around thanks to Salvaña and his colleagues, who kicked off a new HIV/AIDS awareness campaign with a 5 September benefit concert in Manila. They are following up with school visits and plans to pressure the government to make HIV/AIDS intervention a priority. "This is our window of opportunity," says Salvaña, "if it hasn't passed yet." **–DENNIS NORMILE**

CREDITS: (TOP TO BOTTOM) PHILIPPINE RED PARTY; (SOURCE) FRITS VAN GRIENSVEN/PHILIPPINES DEPARTMENT OF HEALTH

Enforcer. Carlos Minc at a government raid in June near Santarém, Brazil, with illegal timber.

CREDIT: JEFFERSON RUDY

NEWSMAKER INTERVIEW

Carlos Minc Pushes a Bold Agenda As Brazil's Environment Chief

BRASÍLIA—Carlos Minc has a reputation for bold strokes. During Brazil's military dictatorship, he was accused of helping an armed leftist group make off with a 350-kg safe belonging to a corrupt politician. The reputed haul: $2.5 million. Minc was tortured and jailed and spent years in exile. Now Minc, who became environment minister in 2008, is again in the spotlight, this time for his role in engineering Brazil's announcement on 13 November that it would voluntarily make deep cuts in its greenhouse-gas emissions.

As the first major emerging economy to promise cuts, Brazil could seed a diplomatic initiative at December's United Nations summit on climate change in Copenhagen, where countries are attempting to forge a replacement for the expiring Kyoto Protocol. So far, China and the United States, the world's largest emitters, have refused to endorse specific targets. But other countries—including South Korea, which announced a plan for 30% cuts by 2020—have joined Brazil in calling for action.

A 58-year-old geographer rarely seen without his trademark embroidered vest, Minc has a reputation as a media-savvy green leader with pragmatic instincts. He won industry applause when he quickly approved a huge petrochemical plant. But he's successfully pressed for tougher action against deforestation, posing with seized timber or jumping on a tractor to knock down illegal charcoal furnaces. *Science* caught up with Minc in Brasília to ask him how he persuaded Brazil's president, Luiz Inácio Lula da Silva, to adopt voluntary emissions cuts. His answers have been edited for brevity. **–ANTONIO REGALADO**

Antonio Regalado is a writer in São Paulo, Brazil.

Q: Whom do you consider the most influential environmentalist in the world today?
C.M.: I think it's the scientists of the Intergovernmental Panel on Climate Change. The positions that the IPCC took showed people for the first time that global warming is not a fiction and created a global clamor around the biggest planetary crisis civilization has faced.

Q: What is Brazil's mitigation plan?
C.M.: We want to reduce greenhouse emissions by 36 to 39% by 2020, which will actually take us back to where we were in 2003. … A year ago, we didn't have a climate plan or emissions goals, even though Brazil is the fifth or sixth largest emitter.

Q: How did you persuade Brazil's president to agree to the plan?
C.M.: We worked with the scientific community and showed that our mitigation steps, like expanding no-till farming and recuperating degraded land, wouldn't freeze our economy and that they didn't interfere with the creation of jobs or the reduction of hunger and inequality.

Our diplomats worried that if we adopted aggressive limits, we would embarrass our allies, who aren't reducing emissions. India, for instance, is going to double them. But after a lot of discussion with the environment ministers of China, India, and South Africa, they told me they wouldn't be offended if Brazil went ahead, so long as it was clear we were not demanding that they do the same thing. They wanted to make sure our criticisms would be directed more at countries like the U.S., which are major emitters and not doing their homework. So I explained that to the president.

Q: How long did it take to convince the president?
C.M.: Two months. I explained to him that Brazilian society, even the industrial sector, was asking for this, and if we didn't act, it could lead to frustration domestically. Secondly, there was a chance that Copenhagen would be a failure. And Brazil, under no conditions, could associate itself with failure. We had to be a breath of oxygen, of renewable energy, for Copenhagen. So then President da Silva, who is not a scientist but is a great statesman and highly intuitive, sniffed out the best path: that what is good for ecology is also good for him politically, as much internationally as domestically.

Q: How do we know Brazil's mitigation plan is real and not mere talk?
C.M.: Half of our 39% emissions cut will come from the Amazon—reducing deforestation by 80%. We're starting with a baseline of 19,500 square kilometers [of forest lost annually] between 1996 and 2005. This year, we already got to 7500 square kilometers, the lowest level of deforestation in 21 years.

Q: How are you reducing deforestation?
C.M.: I've participated in 28 raids in the Amazon during the last 17 months. When we discovered cattle in environmental reserves, we auctioned them off for the government's antihunger program. I said illegal cattle will turn into barbecue for the poor. People thought I was kidding. We've already donated 20,000 head of cattle to prevent environmental criminals from enriching themselves. The other side is, you need to provide alternatives for sustainable forest management. We established minimum prices for 10 forest products. Just punishing people doesn't work.

Q: Two days after Brazil's announcement, the U.S. and China said no climate treaty was possible at Copenhagen. Your reaction?
C.M.: It was frustrating. I think the citizens of those countries need to speak out forcefully and not accept those positions as inevitable.

Science and the Stimulus

An $18 billion burst of funding is nurturing U.S. research. Where did it go, and is it being spent wisely?

WHAT'S THE BEST WAY FOR THE U.S. GOVERNMENT TO SPEND $18.3 billion on basic research? The question may sound like a scientist's dream. But it's actually an ongoing experiment that began when President Barack Obama signed the American Recovery and Reinvestment Act (ARRA) less than a month after taking office.

The Recovery Act—also known as the stimulus package—allocated $787 billion across the federal government to help lift the country out of the recession. Agencies were told to spend it quickly—before 30 September 2010—and to focus on creating jobs. But the investments are also supposed to help make the country more competitive in the long run.

The huge research windfall—the National Institutes of Health (NIH) was given $10.4 billion, the National Science Foundation (NSF) $3 billion, and the Department of Energy (DOE) received $2 billion for basic research—presented these agencies with a unique opportunity to catch up from previous shortfalls. (Several other agencies received lesser amounts for science, and DOE also received tens of billions for various energy programs.) But it also put pressure on them to find ways to avoid a disruptive boom-and-bust cycle. The 407-page legislation sets out general rules for everyone to follow, with plenty of leeway for individual agencies. This package examines what the three leading science agencies did with that flexibility and how their choices will leave a deep imprint on U.S. research for years to come. The final story in the package examines a controversial area of research jump-started with $1.1 billion in stimulus money: comparative evaluation of medical treatments (p. 1183).

How it happened

Why did the Recovery Act include so much money for science? As a candidate, Obama had talked about the need for an economic stimulus package that included research investments, and his transition team began filling in the details almost immediately after Election Day. Democratic leaders in Congress had their own ideas, dating back to the Innovation Agenda drawn up in 2005 by Representative Nancy Pelosi (D–CA) and others before Democrats won control of Congress and Pelosi became Speaker of the House of Representatives. Both groups sought input from the scientific community, which was happy to provide lists of construction-ready projects and advice on broader areas to fund.

On 15 January, the House put forward its version of the legislation, and its research provisions survived mostly intact after negotiations with the Senate. "We knew there would be competition for the dollars, but we also thought our priorities would compete favorably with any other request because we were all about the future," Pelosi told a friendly audience of university administrators and research advocates last week.

The agencies also did their part. NSF Director Arden Bement, serving a 6-year term that runs until November 2010, had testified repeatedly before Congress that the agency had a $2 billion backlog of highly rated proposals (see p. 1181). Raynard Kington wielded less clout as acting NIH director, but NIH's final tally vastly exceeded what most agency officials had expected thanks to the efforts of Senator Arlen Specter (see p. 1179). DOE's lame-duck political appointees were not involved in any negotiations. But the final bill largely achieved funding levels for science that the department had requested in previous budgets, helping incoming Energy Secretary Steven Chu get off on the right foot. It also launched one of his priorities, a new energy research agency created by Congress that the Bush Administration had opposed (see p. 1177).

Once they received the money, the three science agencies took different approaches to spending it. DOE bankrolled a lot of infrastructure projects, especially at its national labs. NSF chose mostly to fund its backlog of good ideas. NIH did that, too. But it also held two major competitions aimed at pushing scientists in new directions. Despite short turnaround times, the solicitations generated a staggering 26,000 proposals.

Online
sciencemag.org
Podcast interview with author Jeffrey Mervis.

The president's science adviser, John Holdren, was an adviser to the transition team during the negotiations over the Recovery Act and didn't take office until a month after the bill was enacted. But he gives his Administration colleagues a thumbs-up on the decisions they've made. The stimulus money is "addressing pressing national needs" in fields ranging from clean energy and climate change research to science education, Holdren says, as well as responding to the "pent-up demand in the community."

Holdren's predecessor, John Marburger, gives highest marks to DOE's preference for funding construction projects and large instruments. "If you just have a burst of funding," he says, "it makes a lot of sense to use it on capital projects [rather than grants] because it doesn't create the same longtime commitment to people and cause as much disruption when it ends." He thinks that NSF's decision to boost success rates by funding a higher percentage of worthy proposals "was probably the easiest route and a good one." But he worries

Science at work. House Speaker Nancy Pelosi (*far right*) and Representative Rush Holt (*second from left*) helped university officials celebrate a new Web site touting the research achievements of the Recovery Act.

about the long-term impact of NIH's decision to hold new competitions. "NIH has the most difficult problem [after soliciting] a huge amount of proposals and raising expectations," says Marburger, who played no role in negotiating the size of the package. "We'll have to see if they got it right."

Holdren shares that concern about what happens after the stimulus money is spent. He asserts that "agencies are generally doing well in managing their stimulus spending to ensure a smooth transition to the post ARRA-era. … [But] will these strategies be enough to avoid falling off a cliff in the future? Not entirely, simply because the size of the investment makes some dislocations inevitable after the money runs out."

Adopting the same metaphor, Pelosi says that "falling off a cliff is not what we had in mind when we passed" the Recovery Act. And she doesn't think it will happen. "We know that we need to sustain the effort, or increase it, in both the life sciences and the hard sciences," she says. "And no one is more aware of that need than the president."

The White House has made it easy for the public to keep score, with a Web site, recovery.gov, based on data that recipients provide quarterly to their funding agency. A compendium of anecdotes from the grass roots about research projects that address national priorities is available at scienceworksforus.org, a new Web site created by a coalition of universities.

Those snapshots from academia and national labs are intended to sustain support for basic science as Congress races to complete work on the 2010 budget and gears up for what is expected to be a very difficult 2011 budget cycle. Last week, Representative Bart Gordon (D–TN), chair of the House Science and Technology Committee, warned research advocates of the tough road ahead. "The Recovery Act was an anomaly," Gordon explained. "Now we're back into the regular appropriations cycle, and there's a consciousness we've got to get the deficit down." His advice? "We have to get the band back together." **–JEFFREY MERVIS**

Shovel-Ready Science Drives DOE Decisions

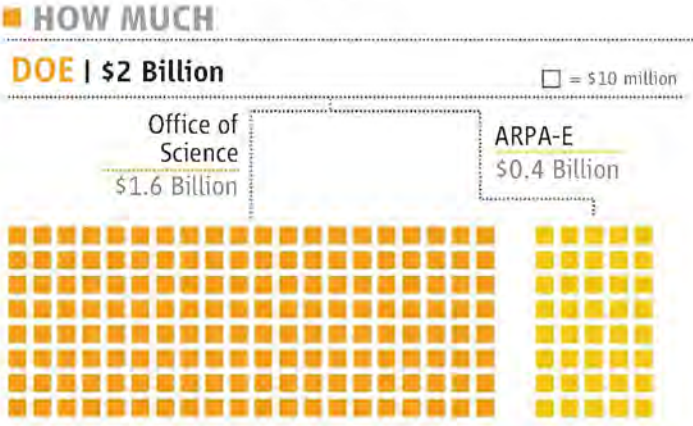
National laboratories enhance their ability to tackle big problems, and ARPA-E gets off to a fast start

In the beginning were the lists.

Candidate Barack Obama had talked about a massive government-spending program to revive the ailing economy, and after the election his transition team moved quickly to flesh out the idea. Sensing an opportunity, a handful of prominent physicists and science lobbyists compiled a list for Congress and the new Administration of all the "shovel-ready" research infrastructure projects at the Department of Energy (DOE)—those approved and ready to go out for bids.

A similar exercise was taking place within DOE itself. Patricia Dehmer, then acting head of the DOE Office of Science, asked the office's six associate directors to give her a list of projects that satisfied two mandates that would appear in the enacted American Recovery and Reinvestment Act: to create jobs, and to provide the scientific and technological base needed for long-term economic recovery. Dehmer added two related criteria. "It had to be high-priority research and construction that was shovel-ready so we could move quickly," explains Dehmer, now back as deputy director since the appointment of William Brinkman to head the science office. "And no, or few, out-year mortgages. I didn't want to reach a cliff in 4 or 5 years."

Two years of tight DOE science budgets had created quite a backlog of approved but unfunded buildings, large instruments, and other research improvements at DOE's 10 national laboratories and various user facilities. The staff-driven list carried a price tag of $5 billion. The



An energy boost. Stimulus funding is accelerating projects in DOE's science pipeline and supporting two rounds of awards at a new research agency.

community's list, which included projects supported by the National Science Foundation (NSF) and the National Institute of Standards and Technology, totaled twice that amount.

Both were put to good use. Congress used the community's list as a guide in deciding to give DOE's science programs $1.6 billion in the final bill. Another $400 million went to the Advanced Research Projects Agency-Energy (ARPA-E), a new research entity housed within

CREDITS (TOP TO BOTTOM): GREG SCHALER/GREGSCHALER.COM; SOURCE: ARRA

the secretary's office. ARPA-E is a favorite of Energy Secretary Steven Chu, who sees it as a way for the country "to hit a few home runs" as it tries to reduce its dependence on foreign oil, curb greenhouse gas emissions, and transition to a low-carbon economy. Even before President Obama signed the legislation on 17 February, however, Dehmer had whittled down her list to fit the expected funding level and obtained approval from Chu, who had been in office for only a few weeks. "We talked through them, and he accepted the recommendations I made," she says.

About three-quarters of the spending was unveiled in March. (Dehmer says a second round of projects, announced in August, needed more time to work its way through the White House approval process. But the belated announcement had nothing to do with the merits of the projects themselves.) Most of the money went to facilities, buildings, and instruments at the various labs, with the expectation that it would be spent before 30 September 2010.

The use of the recovery money to speed up projects already in the agency's queue is expected to pay both scientific and fiscal dividends. Of the $90 million received by the SLAC National Accelerator Laboratory at Stanford University, for example, $33.6 million went to help outfit its newly opened Linac Coherent Light Source (*Science*, 9 October, p. 221). "The LCLS is the future of this lab," says Director Persis Drell. "And having to wait for the instruments to come would have been very frustrating, because we want to get the science going as quickly as possible."

The $65 million that Oak Ridge National Laboratory in Tennessee received for its new chemical and materials sciences building means that hundreds of scientists will be able to move into their new digs sooner, says Associate Laboratory Director Thomas Mason. By going out for bids in a slow economy and with the money in hand, he says, Oak Ridge managed to get an extra 1850 square meters of laboratory space at no additional cost.

Not all of the recovery money is going into infrastructure. And there are even exceptions to Dehmer's rules about spending the money quickly and avoiding mortgages. The biggest involves a program launched in 2008 by Dehmer's former boss, Raymond Orbach, called Energy Frontier Research Centers. Designed to encourage multidisciplinary work on the country's toughest energy challenges, the program was budgeted at $777 million over 5 years. But the Office of Science's 2009 budget was so tight that it could scrape together only $100 million to seed the 46 centers. To accelerate their progress, $277 million in recovery money is being used to fully fund 16 of the university-based centers. The rest will be funded out of DOE's annual budgets for 2010 and beyond. And Dehmer says the Basic Energy Sciences program will have to find money from its budget to pay for any centers it wants to continue after their initial 5-year run.

■ CHIEF ARCHITECT

More than managing. As acting head of science, Patricia Dehmer drew up the funding list and made sure the money got out the door.

■ BENEFICIARY

Building boon. The Recovery Act has meant a $65 million head start for this $95 million chemical and materials sciences facility at Oak Ridge National Laboratory.

DOE is also wading into graduate education for the first time. "We wanted to begin a program in which graduate education was a part of what we did, and we felt that graduate fellowships would attract young people to our portfolio," says Dehmer about the first-time, $12.5 million competition. "But we just didn't have the money to do any more."

Another $85 million of the Recovery Act funds has been earmarked to support early-career scientists at universities and at DOE's own national labs. The new DOE program is modeled after NSF's CAREER awards, with each program office supporting its own cadre of young scientists. "When you have a gold standard, you try to emulate it," says Dehmer. She says DOE hopes that both the graduate fellowships and CAREER awards will reach a "steady-state level" of 400 scientists.

Although DOE's process of allocating its Recovery Act money was opaque to outside scientists, the community seems very satisfied with the results. "I don't know what happened within DOE," says Michael Lubell of the American Physical Society in Washington, D.C., which coordinated the community effort. Once the lists had been compiled, he says, "it was out of our hands. But my sense is that everything that was urgent and that was ready to go got funded. At least I haven't heard anybody complain."

Indeed, directors of several labs that benefited from the recovery money are applauding loudly. "They knew where they wanted their scientific program to be going," says SLAC's Drell. "They do an extremely good job of looking ahead."

Chu knows that Congress is not likely to be as kind to DOE's budget as it was during his first year in office. But armed with the president's promise to double the Office of Science's budget by 2016, he's optimistic that DOE research will not suffer as the Administration attempts to trim overall federal spending.

"It's a flat overall budget," Chu says of rumors already circulating about the president's 2011 budget request. "But [research] is where the federal dollars can go the furthest. And that's why the president is calling for a 10-year doubling of the science budgets despite the fact that, if the DOE budget is flat and you're doubling the Office of Science budget, then you'll have to take it from someplace else [within DOE]. And I'm supportive of that." **–JEFFREY MERVIS**

CREDITS (TOP TO BOTTOM): JASON RICHARDS/OAK RIDGE NATIONAL LABORATORY; COURTESY OF BROOKHAVEN NATIONAL LABORATORY

NIH Hopes Stimulus Isn't a Roller-Coaster Ride

A once-in-a-lifetime increase is a lifesaver for thousands of scientists. But will it also give them a false sense of hope?

Stem cell researcher Michael Kyba hit the jackpot this year. Against daunting odds, Kyba won three research grants funded by the $10.4 billion the National Institutes of Health (NIH) received from the American Recovery and Reinvestment Act. The grants will expand his work at the University of Minnesota, Twin Cities, on how embryonic stem cells become blood cells. He's also leading a new $2 million, multi-institution collaboration on using stem cells to study and treat muscular dystrophy. Taken together, the awards will double the size of his lab.

"I absolutely didn't expect all these grants to be funded," says Kyba, who is now scrambling to hire three technicians and three postdocs. But there's a caveat, he points out: "It's only for 2 years."

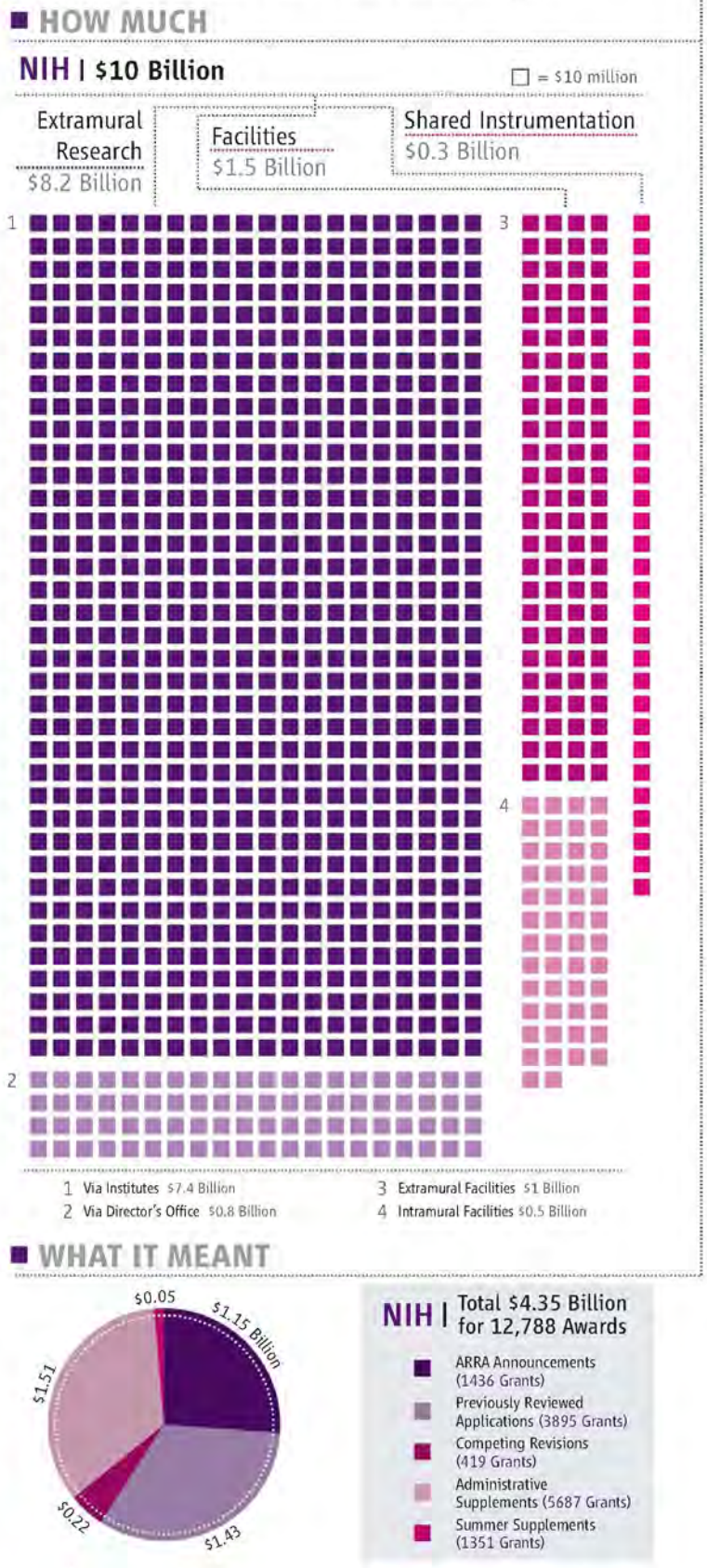
At universities across the United States, and at the NIH campus in Bethesda, Maryland, officials and scientists are taking stock of a once-in-a-lifetime moment in U.S. biomedical research. The stimulus money has kept afloat the labs of thousands of investigators whose proposals had just missed the regular funding cutoff, as well as supporting novel research ideas, new investigators, summer students, and large technology projects. "It's unprecedented, and it couldn't have come at a better time," says Steven Fluharty, vice provost for research at the University of Pennsylvania, which has received $143 million from the NIH windfall.

Raynard Kington, who was interim NIH director from last fall through mid-August, lauds the "group psychological benefit" of increasing NIH's regular budget of $30.5 billion in 2009 by roughly one-third—as well as the "intellectual jolt" that the stimulus provided. "I think it turned out incredibly well," says Kington, once again deputy NIH director since the appointment of Francis Collins as director.

Two aspects of the NIH stimulus funding have attracted particular attention. First, it dwarfs the amount every other federal research agency received. Although some community leaders were concerned initially that it might be too much money to digest—and remain worried that it has set the stage for disaster in 2 years—most have embraced the argument that the increase simply allows the agency to recover from a string of flat budgets. "Steady multiyear budget commitments would be better policy in the long run, but to have not funded NIH during this economic crisis would have been worse public policy," says Elias Zerhouni, who stepped down as NIH director in fall 2008. "We would have lost a lot of young scientists and good research."

Second, NIH chose to hold new competitions for some of the funds rather than drawing from a pool of proposals already in-house, as its sister research agencies have done. That decision triggered a tsunami of applications that led to a projected 1% success rate for the most prominent program, called Challenge Grants. (The actual rate was 4% after NIH institutes quadrupled the initial allocation.) By growing its portfolio, however, NIH also increased pressure on subsequent budgets, because most of those additional grantees are expected to apply for renewals in 2011 and 2012.

The combination of a huge bolus of money and the additional mouths to feed has left many scientists wondering about the real impact of the stimulus spending on the health of the U.S. biomedical research enterprise. "I'm sort of back up in buying power to where I was 5 years ago," says Gerard Evan, a cancer biologist at the University of California, San Francisco, about his $465,000-a-year supplement. "It's great, but the system's still sick."



Spreading the wealth. Extramural scientists and facilities are benefiting from a combination of new awards, supplements, and second chances.

SOURCE (TOP TO BOTTOM): ARRA; NIH (DATA AS OF 26 OCTOBER 2009)

"We gulped"

Although the size of the stimulus may have been a shocker, the idea of giving NIH a big spending boost was not new. In July 2008, senators Tom Harkin (D–IA) and Arlen Specter, the Republican-turned-Democratic senator from Pennsylvania, proposed a 1-year, $5.2 billion increase as part of a 2008 supplemental appropriations bill that was never enacted. That figure, developed with input from NIH, was designed to make up for losses to inflation since 2003, the final year of a 5-year budget doubling. It also provided some extra money for the National Cancer Institute. (Specter has been treated twice for Hodgkin's disease.)

Shortly after the November election, lobbyists and some legislators began talking up a $1.9 billion increase for NIH. President Barack Obama's transition team, which included Collins and former NIH Director Harold Varmus, pushed for more. Specter's staff came up with two numbers—$3.4 billion and $10 billion. The latter number, spread over 2 years, was, like the earlier summer proposal, intended to give NIH the buying power that it had in 2003. In return, Specter wanted NIH to do something novel with the money. Kington says NIH agreed not to just fund more highly rated awards—"We wouldn't simply march down the pay line"—but also to solicit new proposals with short-term research goals that would "get us over some hurdle."

Specter eventually embraced the larger number, and it stuck. The final bill divided the money into $1.8 billion for construction and equipment and $8.2 billion for extramural research. NIH also received $400 million from another agency for research on the comparative effectiveness of medical treatments and procedures (see p. 1183). "We were very pleased. I never for a moment doubted that we would be able to thoughtfully spend the money," says Kington. Another high-level NIH official, however, has a different recollection: "We gulped."

■ BENEFACTOR

Ten-billion-dollar man. Arlen Specter insisted on a monster increase.

A frenzy to apply for the money followed. It was driven by the Challenge Grants—up to $1 million over 2 years—which were developed to address Specter's concern that the money be used for new ideas as well as existing projects. Lured by a 220-page solicitation that covered 15 broad topics, tens of thousands of scientists dropped everything to apply for the $200 million pot. "The strain on our campuses was acute," says Patrick White, vice president for federal relations for the 62-member Association of American Universities.

The total number of Recovery Act applications eventually topped 26,000, far exceeding what NIH receives for its usual thrice-yearly cycle of grant applications in which billions of dollars are at stake. Searching for ways to cope with the onslaught, NIH officials reportedly cracked down on any application that deviated even slightly from the guidelines. Molecular pathologist Anirban Maitra of Johns Hopkins University in Baltimore, Maryland, for example, was part of a Challenge Grant application that NIH rejected because one of the seven investigators exceeded by four the 10 allowable references to publications in his biographical sketch. "We spent so much time putting the grant together. We were sick to our stomachs," Maitra says.

■ BENEFICIARY

A buzz of activity. Michael Kyba's lab is growing thanks to stimulus funding.

What's next

After working overtime all summer, NIH officials managed to disburse $4.35 billion in grants and $379 million in contracts before the 2009 fiscal year ended on 30 September. Because most recipients will receive a comparable amount in 2010, that means NIH has committed 90% of its Recovery money. What remains are some instrumentation and construction awards, plus a small amount for new programs.

Not surprisingly, the money has had a huge impact at universities, say research deans. At Penn, for example, Fluharty says 11 of 23 faculty members on a bridge program funded by the university received NIH stimulus awards. The 1885 grants going to first-time investigators have also meant a big boost for younger faculty members.

Besides these grants to individuals, NIH funded many big-ticket, one-off projects, such as $27 million to create a social network for scientists to find collaborators and share resources. The White House has calculated that more than $1 billion of the $8.2 billion spent on extramural research involves genomics studies, including $175 million for the Cancer Genome Atlas and $64 million to sequence the genes of 8000 participants in long-term heart and lung studies.

The next challenge for investigators with recovery money is to meet NIH's goal of creating or preserving 50,000 jobs. Most universities have a readily available pool of qualified technicians. But the uncertainties of completing their graduate work and defending their doctoral dissertations make it harder to find potential postdocs quickly. Kyba says he will warn candidates that the job might end in 2 years.

University administrators are also very worried about 2011. "We will have built [more] capacity; we will have hired people," says Fluharty. "How will those jobs be maintained?"

NIH is already worrying about the short-term impact of the stimulus on community expectations. Officials anticipate a possible bulge in grant applications in February, fueled in part by the fact that the 12-page limit imposed on Challenge applications will now also apply to standard R01 grants. That equivalency will make recycling rejected proposals a breeze.

Unfortunately for those applicants, however, NIH has already spent its stimulus money. That means Collins must convince Congress that NIH's budget needs to keep growing at least at the rate of inflation. "My sense is that we can make the case by talking about the science, what's being done and what has already been done" on things such as improving survival rates for diseases, he said last week. And although Collins is concerned about the need to boost success rates, he's not planning to use it as his trump card. "I'm not sure it gets the attention of Congress … in the way the science does," he says. **–JOCELYN KAISER**

NSF Boosts Success Rates, But at What Price?

Dipping into a backlog of highly rated proposals is a safe bet. But are there enough high-risk, high-reward ideas in the pot?

Hilairy Hartnett got the bad news in January. The National Science Foundation (NSF) had turned down, for the third time, her application for a CAREER grant, a prestigious 5-year award for tenure-track faculty members who want to combine their research interests with their role as educators. Hartnett's idea to study how the Colorado River takes in and transforms carbon as it flows through the region on its way to Mexico had gotten high marks from reviewers. But her scores were never quite good enough to beat out the very stiff competition.

So Hartnett, a geobiochemist up for tenure this year at Arizona State University, Tempe, decided to rework the idea and resubmit it as a regular proposal to NSF's geosciences directorate. She knew the odds weren't much better, but she had no choice: Researchers can apply only three times for a CAREER award. She also regretted losing the educational component, which would give undergraduates a chance to carry out complex environmental field studies and learn about an important natural resource in their state.

But before Hartnett had done anything, national politics came to her rescue. In May, she got a call from her program manager at NSF, asking if she was still interested in carrying out her CAREER project. Of course, she stammered. But why did it matter? What had changed?

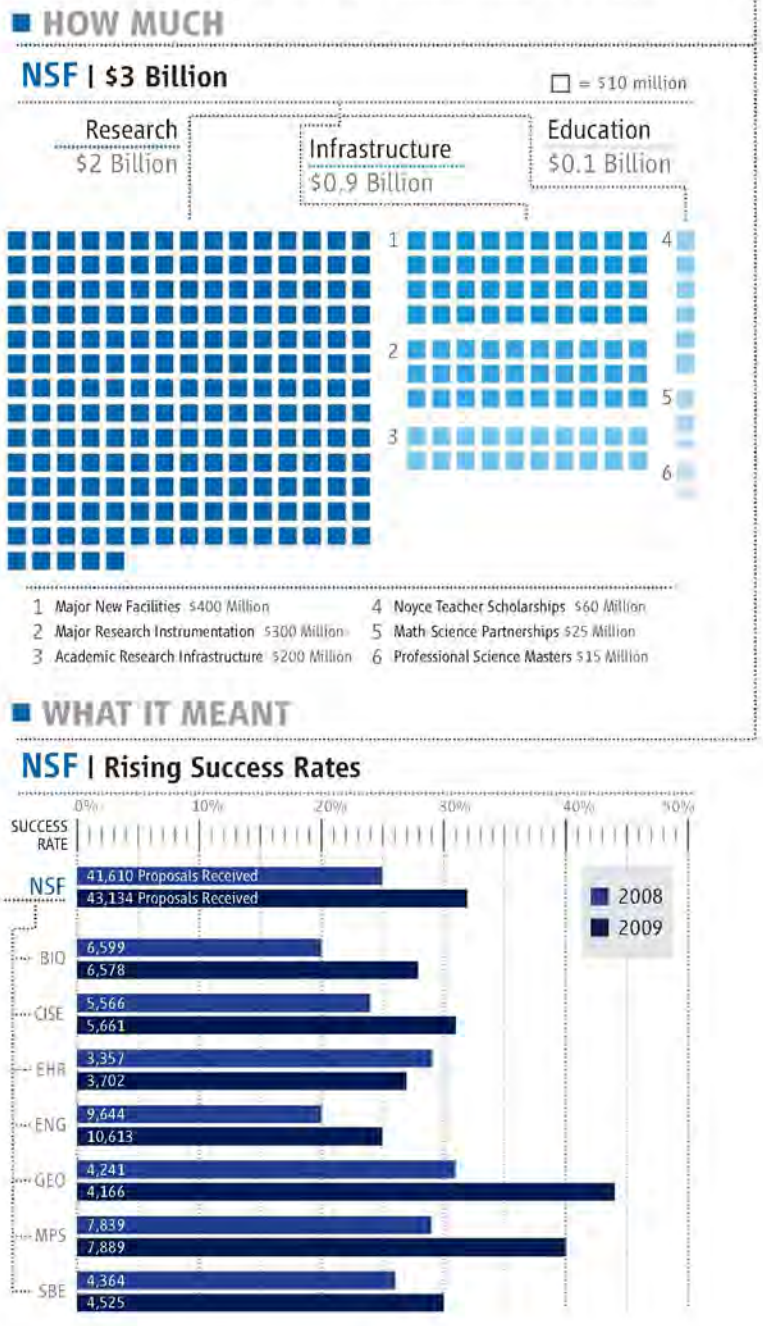
The answer, in short, was the $787 billion American Recovery and Reinvestment Act (ARRA), which was signed into law on 17 February. NSF's slice of the so-called stimulus bill added $3 billion to its regular $6.5 billion budget for 2009: $2 billion for research activities, $900 million for three different infrastructure programs, and $100 million for a handful of specific education and training activities (see top graphic). That huge increase was based on the repeated complaint from NSF Director Arden Bement in congressional testimony that the agency received billions of dollars' worth of good research proposals each year that it couldn't fund. Although Congress specified how NSF should spend the two smaller pots of money, it bowed to NSF's reputation for quality peer review by giving it great leeway in disbursing the largest component.

Bement decided to use the recovery funds to whittle down NSF's huge backlog of highly rated proposals and raise success rates—improving the odds for applicants like Hartnett. Bement also told program officers to give priority to young investigators and to "high-risk, high-reward research." The latter phrase is code for two important policy debates swirling around NSF. "High-risk" addresses concerns from Congress and the scientific community that NSF is too conservative in making its granting decisions. "High-reward" signifies that the agency's portfolio is relevant to important national needs, such as research on climate and energy, that are top priorities for the new Obama Administration.

Bement's priorities were a trifecta of good news for Hartnett. She is a young (albeit at age 41 hardly inexperienced) investigator. Her CAREER proposal was technically still in play. And her project explores the link between terrestrial and aquatic ecosystems, important components in understanding the effects of climate change. "I was stunned, and I'm infinitely grateful," says Hartnett, one of many researchers whose rejection letters were turned into grants once NSF received the recovery money. Coming on the heels of a 5-year, $389,000 NSF grant that expires in March 2010, the 5-year, $573,548 CAREER award will also sustain her lab. "It's dramatically changed my situation."

All told, NSF staffers spent an additional $165 million on the CAREER program, boosting the number of awards by 50%. (Each of NSF's seven directorates participates in the program, with the two largest, engineering and math & physical sciences, making the most awards.) What's more, some 45% of the grants funded with recovery money included at least one first-time grantee, a much higher percentage than normal and another perennial goal of NSF program managers. NSF has also calculated that 10% of the stimulus money (489 awards) has gone to energy-related research, with another 17% (800 awards) supporting climate-related research.

Bement's policies have had a spectacular impact on success rates. The NSF-wide average for competitive awards jumped from 25% in 2008 to 32% in 2009. The geosciences directorate tops the list, with a



Cashing in. Adding $2 billion to a $5 billion research account has given scientists a better chance of winning an award from any one of NSF's seven directorates.

SOURCE (TOP TO BOTTOM): ARRA; NSF

44% success rate (up from 31% in 2008), and the rate for NSF's largest directorate, mathematics & physical sciences, climbed to 40% from 29%. (All numbers are for the fiscal year that ends on 30 September.)

But those higher success rates may not last for long, warn some science lobbyists, if Congress fails to deliver on a promised doubling of NSF's budget over 10 years. "It's happened so many times in the past that the scientific community has come up short," says Michael Lubell of the American Physical Society. Bement says he has faith in Congress and the Administration, pointing to what is likely to be an increase of almost 7% in 2010 and what he hopes will be a healthy 2011 request.

The "summer from hell"

Agencies are required to keep their recovery money separate from their regular appropriations so that the spending can be tracked more easily. But that doesn't mean the funds have to be spent on different programs. Some 307 of the 694 CAREER awards this year, for example, are funded with ARRA money. (That's up from 456 CAREER awards in 2008.) And NSF's engineering directorate was able to boost the number of $2 million grants it made this year for a signature program called Emerging Frontiers in Research and Innovation to 20, from 12 in 2008, by using ARRA money to finance seven of the awards.

Mindful of the rules governing the use of recovery money, Bement decreed that all the projects it supported would be fully funded from the start, rather than receiving the money over several years. But he says that his real focus was on "the fact that we had four and a half months to spend $9.5 billion." Pushing nearly 50% more money out the door before the 30 September end of the 2009 fiscal year translated into 15-hour days for program managers and budget officers across the foundation.

"It was the summer from hell," recalls Jarvis Moyers, head of the division of atmospheric and geophysical sciences within the geosciences directorate. And despite all the extra work, Bement notes that "we didn't get a single nickel to cover our additional administrative costs." One saving grace: NSF's decision not to hold any new competitions (except for two infrastructure programs) meant that the community didn't need to devote additional time to writing grant proposals or serving on review panels.

■ CHIEF ARCHITECT

On the money. Arden Bement made the case for boosting success rates.

NSF's bigger budget has increased political expectations that the agency will place more bets on the type of high-risk, high-reward ideas that have the potential to transform the scientific landscape—and eventually pay off big for the country. One program designed to do exactly that is called EArly-concept Grants for Exploratory Research (EAGER, which until January were called Small Grants for Exploratory Research, or SGER). The grants, which allow investigators to collect preliminary data on a promising idea, are green-lighted by a program officer and don't require external review. Bement has urged all directorates to spend up to 5% of their budgets on such awards.

Program officers traditionally have resisted that suggestion to avoid further depressing already low success rates for their regular programs. And the influx of recovery money doesn't seem to have changed that mindset. Last year, for example, NSF funded 430 SGER awards, representing 0.6% of its research budget. (They are typically smaller than the usual NSF grant.) This year the combined total of the two grants rose to 502. But that's still less than 1% of the budget and far below Bement's target.

■ BENEFICIARY

A fount of knowledge. Hilairy Hartnett shows students how to take a water sample from an acidic hot spring in Yellowstone National Park.

Last month a House of Representatives science panel heard suggestions on how NSF and other federal agencies can do more to foster high-risk, high-return research. Richard McCullough, vice president for research at Carnegie Mellon University in Pittsburgh, Pennsylvania, suggested that NSF create a second program to supplement EAGER "so that, once a transformative discovery occurs, we can accelerate its development." And he pointed out that his university, despite receiving $26 million in stimulus funding and millions more overall in fiscal year 2009 from NSF, has exactly one EAGER award.

Gerald Rubin, vice president for research of the Howard Hughes Medical Institution's Janelia Farms Research Campus in Ashburn, Virginia, wants NSF to go even further. "We fund people, not projects" at Hughes, Rubin explained, predicting that federal agencies would get a much bigger bang for their buck by doing likewise with a small slice of their budgets. As an example, he suggested giving $10 million over 5 years to each of the finalists for NSF's prestigious Alan T. Waterman Prize, awarded to the best U.S. scientist under 35.

Bement doesn't think much of Rubin's suggestion. "As a federal agency, we have to be a bit more prudent," he says. "That's not to say we don't fund people. But sometimes a person who has done outstanding work gets stuck in their own paradigm. … We are always on the lookout for stars, but only those who continue to twinkle."

Holding a new competition, like the one that National Institutes of Health conducted, would have been one way to attract fresh ideas. But that approach had two strikes against it as far as Bement is concerned: It would have depressed success rates and slowed down the process. "We felt that it was important to get as many grants out the door as possible," he says. "So we took full advantage of what we had on hand."

As for whether NSF's overall research portfolio is benefiting society, Jeannette Wing, head of NSF's computer and information science and engineering directorate, has a ready answer. "We funded a project at Stanford University more than a decade ago to create a digital library," she explains. "And on the last page of the final project report, it says, 'and we founded a company called www.google.com.'"

She hardly needs to finish the story. "An investment in one project at one university turned into a completely new industry that transformed society," says Wing. "And you can't predict that."

–JEFFREY MERVIS

CREDITS (TOP TO BOTTOM): ARIEL ANBAR; SANDY SCHAEFFER/NATIONAL SCIENCE FOUNDATION

Medicine Under the Microscope

A billion dollars in stimulus money will go to evaluating which treatments and procedures work best

Long-standing advice to all women in the United States that they should get a mammogram every year after age 40 was overturned last week. After combing through medical data and analyzing six models of disease progression, an expert panel declared that regular mammograms, which expose women to excess radiation and may lead to unnecessary surgery, are more harmful than helpful to someone younger than 50. Even after 50, the panel said, a mammogram every other year would be sufficient. The American Cancer Society and the American College of Radiology immediately rejected the new advice—as did Health and Human Services (HHS) Secretary Kathleen Sebelius the next day—recommending that women follow the familiar old rule.

Such turmoil in the medical ranks could become more common if an evidence-rich kind of analysis backed by the economic stimulus bill takes off. It's known as comparative effectiveness research (CER), and it received a windfall of $1.1 billion in February under the American Recovery and Reinvestment Act. CER employs methods similar to the approach that produced the new advice on mammography.

CER draws data from many sources to reach an evidence-based judgment on the value (or lack of value) of medical techniques and strategies. The process must be rigorous, according to a definition of CER hammered out in June by a federal coordinating group (www.hhs.gov/recovery/programs/cer/draftdefinition.html). Its scope is broad, ranging from comparing drugs in a clinical trial to studying behavior-modification methods to dissecting the impact of health policies. The aim is the same, however: to survey a patient's choices and determine which course works best. The results of a CER study are somewhat like a consumer's guide (see sidebar, p. 1184)—and often as confusing.

Obama Administration officials are enthusiastic about CER and pushed for its inclusion in the stimulus bill as a way to "bend the curve" of health care spending. By this, they mean that it can spot ineffective and overused medical procedures. Armed with such information, they argue, patients and doctors will become more discriminating. In time, they hope, this will slow the growth of U.S. health care spending, which now stands at about $2.4 trillion per year, or one-sixth of the economy.

CER proponents also hope to use the approach to analyze "real world" medical problems. This phrase suggests something different from classic randomized clinical trials, which generally exclude patients who don't fit a specific profile. CER studies, in contrast, may gather records from small clinics and observational studies, taking in people of all ages, including those with complex and overlapping medical problems. Proponents say this could deliver more practical information, more rapidly, than do randomized trials.

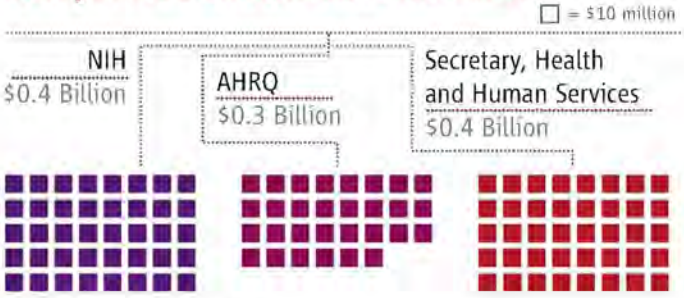
"You can see why we are happy, proud, and excited," says Carolyn Clancy, director of the Agency for Healthcare Research and Quality (AHRQ), which is at the core of the government's push for improving medicine. Her agency has been getting by since 2005 with about $30 million a year for CER, she says. Now the stimulus windfall has "bumped up" the CER portion to $300 million. The National Institutes of Health (NIH) and the secretary of HHS, working with AHRQ, will parcel out the remaining $800 million in stimulus money (see table).

The plan is controversial. Some patient advocacy groups and physicians charge that CER studies could eventually be used to guide insurance and Medicare payments or to overrule a physician's judgment about what is best for the patient. CER has been denounced on blogs as a precursor to "medical rationing."

Key decisions about CER could be affected by the roiling debate.

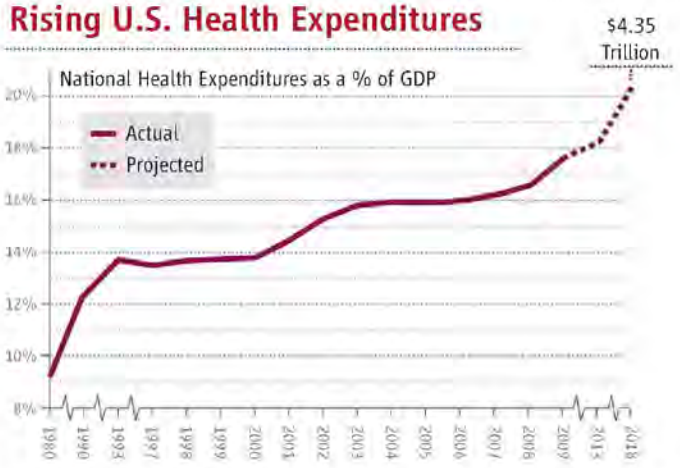
■ HOW MUCH

Comparative Effectiveness Research | $1.1 Billion



■ THE PROBLEM

Rising U.S. Health Expenditures



Relentless. U.S. health costs have been rising steadily and taking up an increasing share of the economy; leaders hope CER can help "bend the curve" downward.

It's not clear, for example, whether the billion-dollar launch of CER in 2009 will be followed by sustained federal support. Nor has it been decided specifically how the government will oversee this research. "There have been open differences about how to frame the research—with some people wanting a more restricted scope and others trying to avoid burdensome restrictions," says Steven Pearson, president of the Institute for Clinical and Economic Review at Massachusetts General Hospital in Boston.

Spending the windfall

With their pockets full of cash, the three agencies with CER stimulus funding began to plunk down commitments this fall. Clancy says the top priority at AHRQ is to pay for "evidence generation," funding investigators to learn what happens to patients with similar ailments who undergo different treatments. Linked to this, Clancy says, is a push to build infrastructure by investing in clinical databases, sharpening analytical methods, and creating networks to disseminate findings.

NIH was first out of the blocks. It has awarded 165 CER grants, totaling $350 million of the $400 million to be spent (projectreporter.nih.gov/files/ARRA-projects.xls). Like others, NIH is funding work on topics from a list of 100 questions drawn up in June by a panel of experts at the National Academies' Institute of Medicine.

Continued on page 1185

SOURCE (TOP TO BOTTOM): ARRA; NIH

A User's Guide to Cancer Treatment

"Is my disease curable?" That's the first question patients usually ask after learning they have prostate cancer, says Patrick Walsh, a surgeon and urology professor at Johns Hopkins Medical Institutions in Baltimore, Maryland. "Then they want to know, 'What are my options for treatment?' " Choose radiation or surgery—and risk severe side effects? Opt for no treatment—and risk having the cancer spread?

Often, there isn't a clear answer. Imaging scans may not accurately identify the amount of cancer in the prostate. Experimental data comparing different treatments are nearly nonexistent. "So you lay out the side effects," Walsh says. "You lay out the fact that you don't know whether surgery or radiation is better. ... It's a very humble conversation."

More than 200,000 men in the United States are diagnosed with the illness every year, making treatment options a prime candidate for review, according to William Lawrence, a medical officer at the U.S. Agency for Healthcare Research and Quality (AHRQ). Lawrence oversaw a comparative effectiveness (CE) review on prostate cancer commissioned by AHRQ in 2005, one of 18 the agency has done.

Timothy Wilt of the Minneapolis VA Medical Center, the review's lead author, and eight colleagues compared the most popular treatment methods, including surgery, radiation, hormone therapy, and "watchful waiting." They also looked at a few not-so-common options, such as cryotherapy, which rapidly freezes and thaws cancer cells to destroy them.

The study, which cost $500,000, followed a standard course laid out by AHRQ, guided by questions from a group of "stakeholders" representing patients, radiation oncologists, urologists, and primary-care physicians. The nine reviewers slogged through mounds of evidence from more than 700 studies, assessed the strengths of doctors and hospitals, weighed side effects, and evaluated outcomes. Data on outcomes and side effects were culled from patient databases, surveys, and clinical trials. They encountered hurdles along the way: Countless definitions of erectile dysfunction and urinary incontinence made side effects tricky to compare. Randomized, placebo-controlled trials were lacking. The team searched for patterns, which they painstakingly graphed as a massive scatter plot, the size of each dot relative to the size of the study.

Data slog. A review team led by Timothy Wilt (seated, *left*) pored over data from more than 700 prostate cancer studies.

Strong differences. An analysis of U.S. prostate cancer treatments by region showed that "watchful waiting" was least practiced in New England.

In February 2008 they delivered their findings: No one treatment is superior. All have adverse effects. Fewer side effects developed among patients treated by surgeons and medical centers that had performed more surgeries. The reviewers also spotted a revealing pattern: Doctors are more likely to recommend the treatment they practice. Wilt says: "Surgeons recommend surgery, radiation oncologists recommend radiation therapy, and very few individuals recommend watchful waiting, even for men with relatively slow-growing tumors, despite the fact that evidence suggests that all three treatment options are quite acceptable. Both patients and providers would be better served by knowing this information and having it displayed and available to them."

Opinions of the review are mixed. Some say it provides important guidance to people struggling with complex decisions. AHRQ produced a pair of guides summarizing the review for doctors and patients. More than 15,000 people ordered the guides for the prostate cancer study, according to AHRQ. "Often, what happens when you get diagnosed with something is you're kind of shell-shocked and you want to read up," says Jean Slutsky, director of AHRQ's Center for Outcomes and Evidence. These guides, she says, help patients organize their thoughts and understand the research.

Others say the prostate cancer review illustrates the limitations of CE research. Walsh, who has published his own book, *Dr. Patrick Walsh's Guide to Surviving Prostate Cancer*, says the AHRQ clinicians' guide provides information that a doctor should already know, and the consumers' guide is full of questions: "They did their best, but I don't find it terribly helpful in the treatment of my patients."

Sean Tunis, director of the Center for Medical Technology Policy in Baltimore and an expert on CE research, agrees that the review fails to address some essential questions. It represents something that occurs commonly in CE research, Tunis says: "A lot of times people mistakenly think that you can do a systematic review of what's known, and you'll come to a conclusion on what's more effective or less effective, more expensive or less expensive. But most come to the conclusions this one did: The studies we'd like to have haven't been done."

Studies to fill research gaps need to be funded faster and tackled more efficiently, Lawrence says: "That's one area where I think you'll see more from us in the future."

–JENNY MARDER

Jenny Marder is a writer in Washington, D.C.

CREDITS (TOP TO BOTTOM): APRIL EILERS/VA MEDICAL CENTER, MINNEAPOLIS

Continued from page 1183

(www.iom.edu/en/Reports/2009/ComparativeEffectivenessResearchPriorities.aspx). NIH Director Francis Collins noted in October that much of this research is familiar: "Eighty-eight of the 100 priorities in the plan are already covered" by NIH-funded investigators, he said in a talk at AAAS (which publishes *Science*).

Typical of the NIH winners is a $2 million project that will zero in on the use of advanced medical imaging, a controversial topic. New machines are being deployed rapidly, says Bruce Hillner, a physician-researcher at Virginia Commonwealth University in Richmond who's involved in the project. X-ray computed tomography, magnetic resonance imaging, and positron emission tomography (PET) are now used for many purposes, from checking out nonspecific "belly pain" to detailed monitoring of cancer as it progresses or recedes. "For more than a decade, the use of these technologies has increased 15 to 20% every year," Hillner says. Studies have shown that specific uses have great value, but indiscriminate use may have penalties, such as triggering "futile surgery"—life-threatening operations that lead to no benefit. Furthermore, the imaging boom isn't driven strictly by medical evidence; it may also be fueled by "self-referral" by physicians who have a stake in imaging services.

Worth it? An important question for comparative effectiveness research is whether medical scans are used appropriately.

Hillner is co–principal investigator at a repository, the National Oncologic PET Registry (NOPR), that has gathered data from U.S. cancer-related PET scans since 2006. He's teaming up with Dartmouth Medical School researchers William Black and Anna Tosteson, who lead a new collaboration to review cancer imaging. Hillner says he plans to "work backwards" from the NOPR database, which was designed for this purpose, to see if physicians followed the cancer management plan they described when they sought reimbursement for a PET scan. Among the questions he'll be asking are: Did the plan change? What happened to the patient? How did PET affect surgery decisions, including futile surgeries? Did physicians generally use multiple overlapping technologies, or did they use only PET?

Like NIH, which is putting nearly all of its money into research, AHRQ will invest $100 million of its $300 million in investigator-initiated clinical studies called CHOICE awards. It plans to make 10 of these at roughly $10 million apiece, running for 3 years each. For these, the White House has waived the rule that stimulus money be spent within 2 years. "They are big grants" for the field, says Tosteson.

Evidence, please. Carolyn Clancy directs AHRQ, which funds research on what works in medicine.

CREDITS (TOP TO BOTTOM): © ROBERT LLEWELLYN/CORBIS; AHRQ

AHRQ hasn't announced the CHOICE winners as of yet. But its detailed spending plan, required under the Recovery Act, includes $48 million for patient registries, $25 million for evidence synthesis, $29.5 million for dissemination of results, and $20 million for training. (HHS hasn't released a list of grants or a spending plan.)

How will findings be used?

The plan to boost CER ran into flak early this year when opponents said studies might be used to cut back on specific treatments in an attempt to lower costs and ration medical care. Critics included well-known physicians, patient advocacy groups, and members of Congress.

One surprising blast came from the respected hematologist and *The New Yorker* writer Jerome Groopman and his wife, Pamela Hartzband, both of Harvard Medical School in Boston. Their 31 August editorial in *The Wall Street Journal* suggested that a national effort to define the best practices in medicine would end up describing only averages and not what's good for a specific patient. They warned that economic incentives intended to nudge doctors toward following national rules—as some would like to do—would mean that "federal bureaucrats are directing health decisions."

Advocacy groups have been outspoken, too. Tony Coelho, a former Democratic congressman from California who heads the new Partnership to Improve Patient Care, has kept up a steady drumbeat of warnings about the need to circumscribe CER. The group argues that CER is fine if used to inform patients and doctors, but it should not be used for "making centralized coverage and payment decisions or recommendations." Coelho's group is supported by the Biotechnology Industry Organization, the Pharmaceutical Research and Manufacturers of America, and the American Association of Neurological Surgeons, among others.

In response to such concerns, Congress decreed in a conference report on the stimulus bill that CER should not examine cost issues. And members of Congress specifically banned using such research to "mandate" changes in Medicare. The future use of CER hangs in part on whether this ban is continued. "Even without cost-effectiveness information," says Harvard historian of science Jeremy Greene, an expert in this subject, "I think CE information is essential to improving the quality of our health care system."

The health care reform bills now before Congress may also determine how CER is managed. The House of Representatives–passed bill and the version drafted by Senate Democrats both seek to reduce friction over CER findings by ensuring that the government consults with stakeholders—patients, doctors, and industry representatives—before making big investments in CER or acting on research findings. But the House and Senate proposals differ on who calls the shots. The House plan would rely essentially on AHRQ to set research policies, with input from advisers. The Senate approach would vest authority in an independent, nongovernment corporation.

Congress is treating the subject with caution precisely because CER promises a radical change in the way medical practices are evaluated, bringing a lot of new data to bear on decisions. And as the brouhaha over screening for breast cancer shows, new evidence isn't necessarily welcome. **–ELIOT MARSHALL**

LETTERS

edited by Jennifer Sills

Retraction

WE WISH TO RETRACT OUR REPORT (*1*) IN WHICH we report that β–N-acetylglucosamine-serine can be biosynthetically incorporated at a defined site in myoglobin in *Escherichia coli*. Regrettably, through no fault of the authors, the lab notebooks are no longer available to replicate the original experimental conditions, and we are unable to introduce this amino acid into myoglobin with the information and reagents currently in hand. We note that reagents and conditions for the incorporation of more than 50 amino acids described in other published work from the Schultz lab are available upon request.

ZHIWEN ZHANG,[1] JEFF GILDERSLEEVE,[2] YU-YING YANG,[3] RAN XU,[4] JOSEPH A. LOO,[5] SEAN URYU,[6] CHI-HUEY WONG,[7] PETER G. SCHULTZ[7]*

[1]The University of Texas at Austin, Division of Medicinal Chemistry, College of Pharmacy, Austin, TX 78712, USA. [2]Chemical Biology Section, National Cancer Institute, Frederick, MD 21702, USA. [3]Rockefeller University, New York, NY 10065, USA. [4]6330 Buffalo Speedway, Houston, TX 77005, USA. [5]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095–1569, USA. [6]University of California, San Diego, CA 92121, USA. [7]The Scripps Research Institute, La Jolla, CA 92037, USA.

*To whom correspondence should be addressed. E-mail: schultz@scripps.edu

Reference

1. Z. Zhang *et al.*, *Science* **303**, 371 (2004).

Where Are the Parasites?

THE REVIEW BY E. POST *ET AL.* ("ECOLOGICAL dynamics across the Arctic associated with recent climate change," 11 September, p. 1355) paid little heed to parasites and other pathogens. The rapidly growing literature on parasites in arctic and subarctic ecosystems provides empirical and observational evidence that climate-linked changes have already occurred. The life cycle of the protostrongylid lungworm of muskoxen, *Umingmakstrongylus pallikuukensis*, has changed (*1*), and the range of that organism and the winter tick, *Dermacentor albipictus*, has expanded (*2*).

On the decline. Pathogens may play a role in the decline of caribou populations across the North.

CREDIT: SUSAN J. KUTZ

Extremes in temperature and the hydrological cycle, predicted in most climate scenarios, can result in epidemic disease outbreaks in arctic-adapted species such as reindeer and muskoxen, with substantial economic costs for northern aboriginal peoples (*3*–*7*). Similarly, increased frequency and magnitude of flooding might enhance transmission of waterborne pathogens such as zoonotic strains of *Giardia*, in and between terrestrial and marine systems (*8*, *9*).

Parasites whose stages in the environment are buffered by gastropod or insect intermediate hosts/vectors have the potential to increase in abundance and distribution, whereas for those with life stages that develop freely in the environment, extreme variability in microhabitat temperatures and humidity might either increase or reduce their abundance (*2*, *10*).

Given the low species diversity of arctic ecosystems, and the potentially reduced immunocompetence of arctic species (*11*), these host systems may be particularly sensitive to parasitic invasions (*2*). Invasions will occur primarily through range expansion of

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

more southerly host species, through ongoing wildlife translocations, and increasing pressures for domestic animal agriculture. All will radically alter the existing parasite fauna and lead to parasite-mediated competition between current residents and newly arrived host species. This might in turn lead to the loss of parasite diversity as arctic-adapted hosts and their endemic parasite species become increasingly displaced by competitive interactions. Such changes will have profound consequences for ecosystem structure and function and directly impact the health, economy, food safety, food security, and cultural activities of northern peoples.

SUSAN J. KUTZ,[1]* ANDY P. DOBSON,[2] ERIC P. HOBERG[3]

[1]Faculty of Veterinary Medicine, University of Calgary, Calgary, AB T2N 4N1, Canada. [2]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [3]U.S. National Parasite Collection, U.S. Department of Agriculture, Agricultural Research Service, Beltsville, MD 20705, USA.

*To whom correspondence should be addressed. E-mail: skutz@ucalgary.ca

References

1. S. Kutz, E. P. Hoberg, L. Polley, E. Jenkins, *Proc. R. Soc. B Biol. Sci.* **22**, 1581 (2005).
2. S. J. Kutz *et al.*, *Vet. Parasitol.* **163**, 217 (2009).
3. S. Laaksonen, M. Solismaa, R. Kortet, J. Kuusela, A. Oksanen, *Parasit. Vectors* **2**, 3 (2009).
4. S. Laaksonen, J. Kuusela, S. Nikander, M. Nylund, A. Oksanen, *Vet. Rec.* **160**, 835 (2007).
5. K. Handeland, T. Slettbakk, *J. Vet. Med. Ser. B* **41**, 407 (1994).
6. J. E. Blake, B. D. McLean, A. Gunn, *J. Wildl. Dis.* **27**, 527 (1991).
7. B. Ytrehus, T. Bretten, B. Bergsjo, K. Isaksen, *EcoHealth* **5**, 213 (2008).
8. S. J. Kutz *et al.*, *Parasit. Vectors* **1**, 32 (2008).
9. A. J. Parkinson, J. C. Butler, *Int. J. Circumpolar Health* **64**, 478 (2005).
10. S. J. Kutz, E. P. Hoberg, J. Nagy, L. Polley, B. Elkin, *Int. Comp. Biol.* **44**, 109 (2004).
11. T. Piersma, *Oikos* **80**, 623 (1997).

Still Vulnerable to Killer Tsunamis

AN UNPRECEDENTED RETREAT OF THE SEA on the morning of 26 December was for some witnesses of the 2004 Indian Ocean tsunami the only warning of the series of deadly waves to follow. Homes were crushed, boats swept away, roads destroyed, coastlines obliterated, and more than a quarter of a million lives were lost (*1*–*4*). Triggered by an underwater earthquake off the coast of Sumatra (Indonesia), the most devastating tsunami in history produced casualties in at least eleven Asian and African countries. In Banda Aceh, the northernmost province in Sumatra, 170,000 perished.

At risk. Many devastated coastal villages, such as Layeun on the northwest coast of Sumatra, are still vulnerable to future tsunamis.

With the help of the international community, the ability to detect tsunamis and provide timely early warnings in the region has improved since 2004. However, on the fifth anniversary of the 2004 Indian Ocean tsunami, many coastal communities are still vulnerable.

Lingering vulnerability stems partly from economics; some of the funds promised for increasing preparedness following the 2004 event failed to materialize as spending priorities shifted in the wake of other disasters worldwide (*5*). However, economics is not the only contributing factor.

Disaster preparedness cannot be accomplished in isolation from the vulnerable communities themselves. The planning, implementation, and maintenance of warning systems require decentralized local initiatives to be coordinated at national levels, albeit with full knowledge, participation, and involvement from local networks. Unfortunately, this is often not the case. Ethnic Moken sea gypsies on Lanta Island in southern Thailand, for example, are ambivalent about the two signal towers and evacuation routes that appeared almost overnight. After all, in 2004 many were racing toward the sea to save their boats after being forewarned of the approaching tsunami by cellphone by family members on outer islands. In addition, many believe that another tsunami of this magnitude won't come again in their lifetime. [Science supports this notion (*6*, *7*).]

Some preparedness measures put in place since 2004 may have limited effectiveness. Newly sanctioned escape routes often parallel low-lying coasts for unrealistic stretches, leaving the elderly, sick, and children at risk to storm surges. Poor or no lighting inhibits visibility at night. As time passes, some signs are beginning to deteriorate; others no longer lead to safe havens. During one early test of the warning signals, some Lanta residents mistook the siren's wail for squawking birds.

Local people may also distrust a national system that appears to largely target high-value tourist coastlines. Throughout the region, many less-developed yet highly populated coasts—Banda Aceh included—are still not protected sufficiently by early warning systems. Several more years will likely lapse before a fully operational system covering the entire Indian Ocean can be established. For some low-lying coasts adjacent to the fault line that triggered the 2004 tsunami, even instantaneous warnings may fail to prevent tragedy.

Socioeconomic conditions and population pressures force many to remain living in harm's way. Communities demolished in 2004 were largely rebuilt in place, where they will always be at risk. Some title-less land holders have been evicted from coastal buffer zones, but in other instances valuable squatter lands have been snatched away by the affluent or entrepreneurs whose actions will likely transfer the risk to richer sectors (*8*). With nowhere to go, the landless are returning to the coasts—this time in potentially more vulnerable areas.

We need a new paradigm for sustainable coastal zone management that recognizes early warning alone is not sufficient for saving lives. Strengthening the livelihoods of coastal communities, disaster-proofing public spaces, relocating vulnerable populations, and preserving coastal ecosystems will go a long way in reducing vulnerability in developing areas. However, local leadership in conjunction with community-based organizations that bridge gender, ethnic, and religious lines must ultimately embrace cautionary principles that are not reliant on top-down measures requiring substantial external aid. The recent 2009 tsunami in Tonga and the Samoas demonstrate that basic awareness of how to behave during a tsunami can help reduce loss of life.

A. D. ZIEGLER,* P. P. WONG, C. GRUNDY-WARR

Department of Geography, National University of Singapore, 117570, Singapore.

*To whom correspondence should be addressed. E-mail: geoadz@nus.edu.sg

CREDIT: WONG POH POH

References and Notes

1. R. Cochard *et al.*, *Perspect. Plant Ecol. Evol. Syst.* **10**, 4 (2008).
2. K. Hirata *et al.*, *Earth Planets Space* **58**, 195 (2006).
3. J. Rigg *et al.*, *Geog. J.* **171**, 374 (2005).
4. Asian Disaster Preparedness Center, *Regional Analysis of Socio-Economic Impacts of the December 2004 Earthquake and Indian Ocean Tsunami* (ADPC, Bangkok, 2006).
5. I. Kellman, *Disast. Prev. Manage.* **15**, 178 (2006).
6. K. Jankaew *et al.*, *Nature* **455**, 1228 (2008).
7. K. Monecke *et al.*, *Nature* **455**, 1232 (2008).
8. P. P. Wong, *Ocean Coast. Manage.* **52**, 405 (2009).
9. We acknowledge J. Yuen, N. Adbul-Jamal, K. Lek, and S. H. Ho (NUS Singapore) for help with fieldwork.

Teaching Strategies Based on Research

THE LETTER "TEACHING AND LEARNING STRATEGIES that work" (R. Hoffman and S. Y. McGuire, 4 September, p. 1203) highlights how scientists—or, at least, these chemists—are isolated from "education." In the Letter, two chemistry professors described educational strategies they have "developed" over the course of their 40 years of combined teaching experience. We were struck by the fact that they made no obvious reference to the discipline of education. In the past 40 years, there have been thousands of research studies documented in the education literature. These research studies detail and inform us of the successes and failures of teaching strategies and their impact on the learner at all stages of development.

Colleges of Education within universities are grounded in research and a body of literature on teaching and learning. All professors wishing to teach should be informed by past research in education, just as we expect chemists to be informed by past research in chemistry. Why is it that professors who earn doctoral degrees in other disciplines automatically presume to be teachers?

Unfortunately, this approach is rampant in academia and even in this journal's sister publication, *Science Signaling*. Upon review of the 57 abstracts (going back to 2003) that *Science Signaling* lists as "Teaching Resources," not one was written by an educator. We will see improvements in teaching sooner if we are informed by research on teaching strategies proven to maximize learning across all developmental levels.

P. DEE BOERSMA[1]* AND LIN ZURFLUH[2]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA. [2]Seattle, WA 98102, USA.

*To whom correspondence should be addressed. E-mail: boersma@u.washington.edu

CORRECTIONS AND CLARIFICATIONS

News of the Week: "When counting jobs isn't enough" by J. Mervis (6 November, p. 783). Susan Sedwick is associate vice president for research and director of the office of sponsored projects at the University of Texas at Austin.

News of the Week: "Beyond Thailand: Making sense of a qualified AIDS vaccine 'success'" by J. Cohen (30 October, p. 652). The story overstated IAVI president Seth Berkley's thoughts about the future of the two vaccines used in the study. Berkley did not dismiss the possibility of other vaccine developers attempting to improve the combination of these two vaccines used in the trial and continuing with their development. In addition, the table on page 653 should have noted that it was "%" Incidence/Yr.

News of the Week: "Research wolves of Yellowstone killed in hunt " by V. Morell (23 October, p. 506). The name of a scientist and his university were spelled incorrectly. The correct spellings are Daniel MacNulty at Michigan Technological University.

News Focus: "Race for the heavens" by Y. Bhattacharjee (23 October, p. 512). The Hale Telescope is located on the Palomar Mountain, not Mount Palomar. Also, Richard Ellis is a professor at the California Institute of Technology, not the University of California, Berkeley.

SOCIOLOGY

Asking About Origins

Gagan D. S. Sood

Without any fanfare, John Levi Martin throws down the gauntlet in the opening sentence of *Social Structures*: "We set out, then, to determine where structures … 'come from.'" This tersely summarizes the overarching research agenda motivating the book. The social structures to which Martin gives prominence embrace the smallest and the simplest ("cliques" and "hubs," in the jargon of sociologists) as well as the very big ones on which the modern state is founded (most notably, the civil service, the political party, and the command army). As for the whence in his opening sentence, Martin (a sociologist at the University of Chicago) is referring both to the nature of these structures—their form and content—and to the ways in which they emerge, dissolve, or are transformed.

Social Structures

by John Levi Martin

Princeton University Press, Princeton, NJ, 2009. 408 pp. $39.50, £27.95. ISBN 9780691127118.

A scholarly interest in such matters is quite natural and readily justified. Throughout history, structures have constrained and enabled many features of human social life. If we wish to comprehend with any rigor the general lineaments of and possibilities for the societies that populate our contemporary world, we must grapple with the why, where, what, and how of their component structures. Scholars in the West have been trying to do this for at least two centuries, ever since the idea of society was intellectually unshackled from divine and kingly power. Martin's ongoing research harks back to the genesis of this tradition in seeking to address comprehensively the social structures that loom so large in our imaginations and in our experiences of daily life.

While his interest in structures has a venerable lineage, Martin's approach is highly distinctive. He critiques and contrasts it to functionalism (which pivots on why structures exist), extreme Rankean historicism (which holds that context is effectively the be-all and end-all in thinking about the past), and erudite, though too often sterile, social theorizing. In their place, he offers an approach that (at least as implemented) is structural, processual, and historical. It focuses attention on the substance of individual relationships, the structures associated with them, and their development and subsequent fate. It conjoins analysis of their objective patternings to subjective, or cultural, understandings of their observed regularities. And Martin applies it in a mode that is at times dialectical, at times inductive, striving throughout (not always successfully) to remain within touch of empirically verifiable reality.

His definition of structure in all this owes a considerable debt, one freely acknowledged, to the scholarship on social networks pioneered by Harrison C. White. In his resolute determination to yoke structures, whatever their size and complexity, to the specific interactions of individual persons, Martin has taken to heart Georg Simmel's influential writings on the general nature of society. But in the final reckoning, he stands on his own, for much of the book a knight's move away from received scholarship.

Putting aside Martin's self-categorization of his approach and reasoning, let's consider this work on its own merits. What we find is that, despite the frequently invoked powerful framing device, this is not a seamless monograph. Rather, it is a book of two halves (separated by the fifth chapter, which occupies a class of its own). Although there is no denying the meaningful links between these halves, I found the space devoted to the elaboration of either (especially the first) is not justified in view of the other. So, it appears sensible to treat each part separately.

CREDIT: *HUMAN PYRAMID*/BRAND X PICTURES/GETTY IMAGES

The first half of the book draws on thought experiments and findings from sociological, anthropological, zoological, and epidemiological studies. Martin uses this evidence to examine in detail three types of social relationships—which he glosses as "symmetric," "asymmetric," and "antisymmetric," noting that his conception of these terms is "somewhat idiosyncratic"—and the kinds of simple structures that might arise spontaneously from them. In particular, he explores the local, pure structures (constituted solely from one of these three types) that these relationships have an "inherent tendency" to produce. Among the examples discussed are circular generalized exchanges, which may derive from asymmetric relationships of gift-giving, and "pecking orders," which can be produced by antisymmetric relationships of dominance. In each case, Martin discovers that either such structures do not exist in reality or, if they do, they tend to be fragile, able to survive only in a narrow range of environments. More important, however, such structures (with very few exceptions) do not concatenate to form larger structures. The structures that do exist tend to be more complex and are often not clear-cut, their constituent relationships differing in one or more of their attributes from the three elemental relationships. A telling instance of this is furnished by the small world–like, "clumpy" networks generated by relationships of acquaintance. Although close in nature to symmetric relationships, these prove on inspection not to be inherently symmetric: "There is both more and less equality in such relationships than we assume."

Martin argues that the main reason for the general incapacity of simple structures to concatenate while maintaining the integrity of their founding relationships is the excessive and unrealistic demands that would be placed on individuals. There are, however, exceptions—perhaps the most crucial being antisymmetric patronage relationships. And it is these that lie at the heart of the second half of the book. Here, Martin carefully develops a set of variegated and sensitive, but ultimately stylized, historical narratives. In contrast to the early chapters, these are largely substantiated by recourse to historical evidence (most of it pertaining to classical Greece and Rome, medieval Europe, and the United States).

Beginning with patron-client relationships, he first shows how these might com-

The reviewer is at the Department of History and Civilisation, European University Institute, Villa Schifanoia, Via Boccaccio 121, I-50133 Firenze, Italy. E-mail: gsood@eui.eu

bine to form patronage triangles, which, in turn, tend to concatenate into pyramids. He then considers what accounts for their seeming prevalence and stability in the past. Martin maintains that these structural forms "thrive in conditions of inequality" and arise "when there is a conscious attempt to ameliorate [the preexisting] inequality by the introduction of diffuse personal relationships." But times change and with them structuring principles. This can result—as occurred with the rise of the modern state in Europe—in the breakdown and transformation of patronage structures, a process that Martin sees as of great historical importance. Technically, what occurred is that mediation (which he argues is a core attribute of these structures) was rearticulated through the imposition of transitivity. This process converted the pyramidal patronage structures into "command tree" structures, which are far more effective in facilitating the kind of collective action required by modern states. Martin elucidates these processes (emphasized throughout as complex and seldom unidirectional) by following the historical emergence and evolution of the army in feudal and postfeudal Europe (and today exemplified by the U.S. military) and of national political parties in the United States. To these, Martin adds a brief survey of the establishment of what he (with many professional historians) claims forms the third central pillar of the modern state, the civil service. Although differences abound, his analysis highlights the remarkable purported similarities and parallels in the stories of the development of each of these structures.

The book is without doubt an eclectic, ambitious, provocative, sophisticated, and instructive undertaking. Much of its substance (which has merely been hinted at above) can be viewed—and assessed—in terms of its three potentially most important contributions. First, it brims with many strong (and often plausible) arguments about the historical social origins of the modern state. The evidentiary basis for these arguments, however, is largely confined to the European (and American) experience. In generalizing from this, Martin lays himself open to charges of unwarranted and arguably pernicious Eurocentrism. Furthermore, in the prosecution of his case studies the author needed to engage more fully with the contested nature within serious historical scholarship of many of the staging posts in his narrative accounts. Second, Martin's investigation of structures and their genealogies deepens our understanding of the guiding principles in social life. Here too he has a great deal to say that is interesting and appears to ring true. But this is heavily compromised by weaknesses in the threads joining his discussions of local and big structures. Third, the book's general approach in itself may come to be viewed as a major advance. It operates on a middle level between theory and data, and it offers a realistic prospect of extending our metaphysical notions of social structures and, thus, of the society in which they are observed. The approach as expressed in the book, however, should be regarded at best as transitional. We need more consideration of the basic methodological conundrum of how to elucidate a structure's genealogy if the processes that produce the structures tend to efface their genealogies. This issue throws into sharp relief Martin's problematic conceptualizations of historical and analytical interpretations, induction, and the commensurability of fundamentally different types of evidence.

It is too early to tell whether Martin's explanation of where structures come from will stand. "But if we are to see whether structural analysis can make a contribution to a general sociology, even simple initial accounts are encouraging." From this vantage, there are many reasons to be encouraged by—and to applaud—this work. The value of the book and of its larger research agenda might, thus, lie not in having produced clear answers to the questions posed at its outset but in suggesting powerful and promising ways in which that fundamental topic might be approached. For this alone, *Social Structures* deserves a wide readership and its ideas a sympathetic hearing.

10.1126/science.1181550

PHYSICS

Quantum Theory's Silent Pioneer

Mary Jo Nye

Graham Farmelo's *The Strangest Man* is a consummate and seamless biography of the brilliant and eccentric physicist Paul Adrien Maurice Dirac (1902–1984), whom theoretical physicists often rank as second only to Einstein in originality and brilliance. In 1963, Robert Oppenheimer wrote Dirac that his photograph would be mounted at the Institute for Advanced Study in Princeton next to a likeness of Einstein: "You two are alone on that wall." Dirac is revered in Great Britain at Westminster Abbey, where his commemorative stone rests next to Isaac Newton's gravestone and a few steps from Charles Darwin's burial place. Yet, unlike Einstein, Newton, and Darwin, Dirac is not familiar to the general public.

This is partly Dirac's own doing. Dirac avoided journalists and the limelight. He considered declining the Nobel Prize rather than face publicity and attention. He refused a knighthood, although he later accepted the Order of Merit, which does not confer a title. As the Lucasian Professor of Mathematics in Cambridge, he developed a good lecturing style, but he sometimes paused in midsentence—even for as long as a minute—to carefully choose the next word. He traveled widely for conferences and visiting appointments. He loved walking and gardening. He had a few close friends, among them the boisterous Russian physicist Peter Kapitza and the quieter English physicist Patrick Blackett, whose left-wing political views he shared. On most social occasions, including high table at St. John's College, Cambridge, Dirac was superhumanly quiet and uninterested in his surroundings. Around 1925, colleagues invented the "dirac" unit for the smallest imaginable number of words that someone could utter in company: an average of one word an hour. Much later, Niels Bohr said of Dirac that he was the "strangest man" who ever visited Bohr's theoretical physics institute in Copenhagen.

This is the first biography of Dirac since Helge Kragh's of 1990 (*1*). That focused on Dirac's scientific work while giving some attention to details of his life and to the philosophical attitudes implicit in his physics. Farmelo (a historian at the Science Museum, London) takes on a great deal more. Making extensive use of private papers still held in the Dirac family, he brilliantly analyzes the

The Strangest Man
The Hidden Life of Paul Dirac, Quantum Genius
by Graham Farmelo
Faber and Faber, London, 2009. 557 pp. £22.50. ISBN 9780571222780.

The Strangest Man
The Hidden Life of Paul Dirac, Mystic of the Atom
Basic Books (Perseus), New York, 2009. 557 pp. $29.95, C$37.95. ISBN 9780465018277.

The reviewer is at the Department of History, Oregon State University, Corvallis, OR 97331–5104, USA. E-mail: nyem@onid.orst.edu

"hidden life" behind Dirac's peculiarities that might help account for Dirac's great success as a theoretical physicist.

The family name originates in southwestern France, where there are towns and villages called Cognac, Cadillac, and Dirac. Dirac's father Charles grew up in Geneva. Charles married a librarian in Bristol and taught modern languages at the Merchant Venturers School. There Paul was an excellent student, especially in the sciences. As Farmelo phrases it, Dirac was something of "an Edwardian geek." Dirac took degrees in electrical engineering and mathematics at the University of Bristol, where he heard the philosopher Charles Broad talk about Einstein's relativity theory. He then won a scholarship to Cambridge, where, under the tutelage of mathematical physicist Ralph Fowler (Ernest Rutherford's son-in-law), he sprang onto the international physics scene. At the time, Cambridge was known for experimental rather than theoretical physics. The German theoretician Max Born was therefore astonished to read Dirac's early work, finding it "one of the greatest surprises of my life." In 1925, Fowler had given Dirac page proofs of Werner Heisenberg's "matrix mechanics paper" (*2*), in which Heisenberg abandoned Bohr's picture of an electron particle orbiting a nucleus and used only empirical data for constructing a quantum theory of the electron. Drawing on his earlier engineering and mathematical studies, Dirac reacted by applying an older mathematical construction from classical mechanics (using what he called "c-numbers" in combination with "q-numbers" of quantum mechanics) to construct a generalized theory of fundamental equations for quantum mechanics. The paper instantly put Dirac in Born and Heisenberg's league.

More surprises were to come. One was the "Dirac equation" (1928), which combined Einstein's special theory of relativity with Erwin Schrödinger's wave equation for the electron. Dirac's equation predicted the existence of the electron mass with precisely the spin and the magnetic field observed by experimenters. In papers of 1929 and 1931, Dirac turned to the troubling implication in his equation that a free electron can have negative energy. Nothing like that had been observed. Why not? Dirac imagined a sea of negative-energy electrons that can only be experimentally detected when one of the electrons jumps away and leaves a "hole" equivalent to a positive charge. The positive hole is a kind of positive antimatter with which the ordinary matter of a negative electron can collide in a process of annihilation that results in emission of radiation.

The reserved and the talkative. Dirac and Feynman at a 1962 conference on relativistic theories of gravitation, Warsaw, Poland.

In 1932, the Caltech cosmic ray physicist Carl Anderson reported the observation of what he thought was a positive electron. Early the next year, Blackett and Giuseppe Occhialini published the results of their systematic experiments demonstrating the existence of the positive electron and the reality of Dirac's annihilation process. Dirac shared the Nobel Prize in Physics in 1933 with Schrödinger. Farmelo further describes Dirac's derivation of the existence of a magnetic monopole and his wartime work in England. He also discusses Dirac's originality in later years, when many people were saying that Dirac had become conservative and unproductive. This work includes Dirac's 1949 suggestion that all physical laws need not show left-right symmetry and time-reversal symmetry, a 1955 proposal that the universe might consist of what Dirac called "strings," and his 1958 prediction of the existence of gravitational energy quanta or "gravitons."

Dirac and Einstein often were compared to one another, and Dirac's classic textbook (*3*) became Einstein's companion for vacation reading. Like Einstein, Dirac was criticized by younger colleagues for clinging to old ideas and prejudices, especially in his unremitting distaste for the renormalization approaches used by Richard Feynman's generation of theoretical physicists to remove infinities from equations. Like Einstein, too, Dirac was a loner in his work habits, an independent thinker, and a devotee to theories that he thought to be elegant. He had a geometrical way of thinking and a mystical faith independent from religious commitment or practice that resulted in his conviction that a beautiful equation has to be correct even if there is no experimental corroboration. Tall and frail, as well as shy and quiet, Dirac looked the part of an aesthete. Digestive problems (finally diagnosed in 1980 as an acid insufficiency in his stomach) led to his painful thinness.

On rare and unusual occasions when he talked of his childhood, Dirac claimed that he never knew love or affection in a family dominated by a bullying father. Charles insisted that Paul sit with him in the dining room at meal times speaking French, while Paul's older brother and younger sister ate with his mother in the kitchen and spoke English. He came to speak little at all, Dirac explained, because of this experience. He also became haunted by memories of the coldness in his late teens of his relationship with his brother, who committed suicide in 1925. At age 35, Dirac married Princeton physicist Eugene Wigner's younger sister Margit (Manci). They had two daughters of their own, and Dirac helped raise her son and daughter from a previous marriage. Margit later recalled that the only occasion she had seen Dirac weep was following news of Einstein's death. Dirac's younger daughter Monica could not remember ever seeing her father laugh.

In Farmelo's view, a possible explanation for Dirac's strangeness lies in autism rather than in environment—especially given the fact that there is conflicting evidence about whether Dirac's stories about the harshness of his childhood are accurate. Farmelo's suggested diagnosis wisely comes toward the conclusion of the biography, rather than serving as a partial framework for the book. The hypothesis adds little to the powerful story that Farmelo tells in this magnificent personal and professional biography, although it does suggest the kind of creativity that may be found in the autistic personality. Dust-jacket praise from playwrights Michael Frayn and Tom Stoppard for Farmelo's reconstruction of Dirac's intellectual adventures and hidden life suggests the dramaturgic potential of this sensitively drawn portrait. *The Strangest Man* might serve as the basis for a drama on the order of Frayn's *Copenhagen*, Stoppard's *Arcadia*, or Ron Howard's film *A Beautiful Mind* based on Sylvia Nasar's biography (*4*) of economist John Nash, who suffered from mental illness. Farmelo has succeeded masterfully in the difficult genre of writing a great scientist's life for a general audience.

References

1. H. Kragh, *Dirac: A Scientific Biography* (Cambridge Univ. Press, Cambridge, 1990).
2. W. Heisenberg, *Z. Phys.* **33**, 879 (1925).
3. P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford Univ. Press, Oxford, 1930).
4. S. Nasar, *A Beautiful Mind: A Biography of John Forbes Nash Jr., Winner of the Nobel Prize in Economics, 1994* (Simon and Schuster, New York, 1998).

10.1126/science.1182289

BIOSECURITY

Educating Scientists About Dual Use

Harmful misuse of biological research must be addressed through education, but such programs are in short supply and can be improved.

Jennifer L. Sta. Ana,[1] Mark S. Frankel,[1] Kavita M. Berger[2]*

In 2004, the U.S. National Research Council (NRC) described the dual use dilemma, the recognition that some legitimate biological research could be misapplied for harmful purposes (e.g., biological weapons) (*1*). The U.S. government's National Science Advisory Board for Biosecurity (NSABB) has released recommendations on communicating, overseeing, and educating about such research (*2*, *3*). The Commission on the Prevention of Weapons of Mass Destruction Proliferation and Terrorism (www.preventwmd.gov/home) recommended that the dual use dilemma and biosafety be taught to all life scientists (*4*). Many other countries have instituted policies to support education (*5*). The UK asks scientists seeking grants if they have considered the dual use implications of their work (*6*, *7*). Biosecurity and dual use research education programs are being developed in Japan, Brazil, and Morocco (*8*). At the 2008 intersessional meetings of the Biological Weapons Convention (BWC), representatives from over a dozen countries acknowledged support for education on biosecurity and biosafety and described education efforts within their nations (*8*). Mandatory education on the dual use dilemma has been supported by the Australian National University (*9*) and is required by law in Israel (*10*). Although these developments represent significant progress, there are still major gaps.

Despite high-level support for education in this area, such resources are few and varied. The University of Bradford and Landau Network Centro-Volta (UB-LNCV) surveyed life science programs in Europe (*11*) and found that only 3 of 57 universities included a biosecurity module and that 22 of 142 degree courses referenced the BWC. In our experience, educational programs that expose life scientists to biosecurity issues are few. The American Association for the Advancement of Science (AAAS) identified 16 programs in the United States that specifically addressed the dual use dilemma (*12*). All programs named by both UB-LNCV and AAAS are included in a publicly available, online repository (*13*).

[1]Program on Scientific Freedom, Responsibility, and Law, American Association for the Advancement of Science (AAAS), Washington, DC 20005, USA. [2]Center for Science, Technology, and Security Policy, AAAS, Washington, DC 20005, USA.

*Author for correspondence. E-mail: kberger@aaas.org

Although these programs reflect modest global interest in educating biological scientists about biosecurity as a component of bioethics, responsible conduct of research (RCR), or biosafety programs, the existing resources and education programs focus narrowly on microbiological experiments identified by the NSABB and the NRC. They fail to address adequately other technologies and nonmicrobial scientific disciplines that could have dual use potential. As technology progresses, scientists will need to be equipped with skills to assess, in an ongoing fashion, risks and benefits of these advances. Tools must be developed to educate all relevant stakeholders (e.g., scientists not trained in the life sciences whose research may impact the biological sciences, institution administrators, and biosafety professionals) about the legal framework under which scientists work and the safety and security risks of research. Instructors and students from undergraduate and high-school levels who work with biological materials should be educated about responsible and safe conduct of research, with programs appropriately tailored for each audience. Topics could include the history of biological weapons, biosecurity national laws and international obligations, codes of conduct and professional norms, ethical considerations of scientists participating in former state-sponsored biological weapons programs, basic ethical decision-making, ways to communicate research to other scientists and the public, and risk-benefit assessment of research.

Another challenge is the lack of funding and time needed for developing and implementing education programs. Our discussions with members of the community indicate that much of the cost of U.S. RCR programs is for salaries of program designers and instructors, which can exceed $100,000. The cost of educating laboratory personnel in biosafety in the United States varies from $4000 to $7000 for researchers and from hundreds of dollars to $4000 for nonscientists, like administrators (*14*). This issue is tied to debates over whether dual use education research should be mandatory or voluntary.

On 16 to 18 November, the U.S. National Academies of Sciences and the Polish Academy of Sciences hosted an international workshop to catalog and assess current programs on the dual use dilemma in all nations (*15*). Many of the issues and efforts raised here were discussed at this international meeting. We encourage the international scientific and security communities to evaluate existing challenges and to propose remedies to ensure that biological research is conducted ethically, safely, and securely worldwide.

References and Notes

1. NRC, *Biotechnology Research in an Age of Terrorism* (National Academies Press, Washington, DC 2004).
2. NSABB, *Proposed Framework for the Oversight of Dual Use Life Sciences Research: Strategies for Minimizing the Potential Misuse of Research Information* [U.S. Department of Health and Human Services (DHHS), Bethesda, MD, 2007]. http://oba.od.nih.gov/biosecurity/biosecurity_documents.html.
3. NSABB, *Strategic Plan for Outreach and Education on Dual Use Research Issues* (DHHS, Bethesda, MD, 2008); http://oba.od.nih.gov/biosecurity/biosecurity_documents.html.
4. Commission on the Prevention of WMD Proliferation and Terrorism, *World at Risk* (Vintage, New York, 2008).
5. Meeting of States Parties, Conference on Disarmament: Refinement, Structure, and Focus, Paris, 1 to 5 December 2008 (United Nations, Geneva, 2008); http://www.unog.ch/80256EE600585943/(httpPages)/008056527905C32EC125755A004B2B1B?OpenDocument.
6. NSABB, Third International Roundtable, Sustaining Progress in the Life Sciences: Strategies for Managing Dual Use Research of Concern, Bethesda, MD, 5 and 6 November 2008 (DHHS, Bethesda, MD, 2008); http://oba.od.nih.gov/biosecurity/biosecurity_documents.html.
7. U.K. Medical Research Council, Position statement on bioterrorism and biomedical research (MRC, London, 2005).
8. S. Miller *et al.*, *Biosecur. Bioterror.* **7**, 93 (2009).
9. Center for Applied Philosophy and Public Affairs, *Ethical and Philosophical Consideration of the Dual-Use Dilemma in the Biological Sciences* (Australian Research Council, Canberra, Australia, 2006); www.ethicsandtechnology.eu/images/uploads/DualUseExecSumm.pdf.
10. Government of Israel, Bill for the Regulation of Research into Biological Disease Agents (2008).
11. UB-LNCV, *Fostering the Biosecurity Norm: Biosecurity Education for the Next Generation of Life Sciences Students*, Como, Italy, 27 October 2008; http://www.centrovolta.it/landau/2008/11/07/FosteringTheBiosecurityNormAnEducationalModuleForLifeSciencesStudents.aspx.
12. J. L. Sta. Ana, M. K. Mohlman, M. S. Frankel, K. M. Berger, *"Professional and Graduate-Level Programs on Dual Use Research and Biosecurity for Scientists Working in the Biological Sciences"*, Washington, DC, 21 November 2008 (AAAS, Washington, DC, 2008); http://cstsp.aaas.org/content.html?contentid=1899.
13. The database, which will continue to be updated, is available at http://cstsp.aaas.org/dualuse.html.
14. J. L. Sta. Ana, K. A. Luke, M. S. Frankel, K. M. Berger, *Biological Safety Training as a Component of Personnel Reliability*, Washington, DC, 17 March 2008 (AAAS, Washington, DC, 2009); http://cstsp.aaas.org/content.html?contentid=2049.
15. U.S. National Academies of Sciences, Education on Dual Use Issues in the Life Sciences, Warsaw, Poland, 16 to 18 November 2009; http://dels.nas.edu/bls/warsaw.
16. Supported by the Department of State and Richard Lounsbery Foundation. We thank those who supplied course material.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1193/DC1

10.1126/science.1176127

CHEMISTRY

Calcite Biocomposites Up Close

Electron tomography is used to probe the processes involved in the formation of biomineral composite structures.

Mark D. Hollingsworth

From echinoderm spines to the exoskeletons of coccolithophores and the prismatic layer structures of mollusk shells, organisms can generate complex and fantastic architectures by depositing minerals such as calcite into matrices containing biomolecules. The biopolymers not only direct the orientation, texture, and morphology of the crystalline network, they are often incorporated within the mineral phase and play integral roles in enhancing mechanical properties (*1*). On page 1244 of this issue, Li *et al.* (*2*) use electron tomography to demonstrate how physical entrapment of biopolymers leads to the formation of high-energy calcite surfaces at the interfaces with the polymer network. The work paves the way toward a clearer understanding of biogenic and synthetic mineral-polymer composites.

Calcite structures in nature. **(Left)** The sea urchin *Paracentrotus lividus* from the Aegean Sea. Each spine is an oriented "single" crystal of calcite. **(Right)** A computer colored scanning electron micrograph of the coccosphere of *Acanthoica acanthifera*.

When thinking of single crystals, the images often conjured up are those of uniformly packed, solid masses bounded by relatively small numbers of symmetry-related facets. It may come as a surprise to know that the exoskeletons of many marine organisms are essentially single crystals of calcite, a mineral that otherwise adopts a rhombohedral morphology with six equivalent faces. Each spine of a sea urchin (see the figure, left), for example, although many centimeters long, is an oriented single crystal whose mosaicity approaches those of solution-grown crystals (*3*).

However, as far as their compositions go, these crystals are anything but single. Biogenic calcite crystals are loaded with biopolymers such as polysaccharides and glycoproteins, which can be oriented along specific crystalline directions within the composite structures (*4*). With certain coccolithophores (see the figure, right), biopolymers are not included to an appreciable extent within the crystalline matrix (*5*). However, a templating process that relies on them is inferred because some kind of chiral template is required to generate the propeller-shaped (chiral) arrangement of the individual coccolith plates, each of which is achiral (*6*).

Although great success has been achieved with optical techniques to ascertain the orientation and distribution of biopolymers within crystalline hosts (*7*), understanding the mechanism of interaction between biopolymers and the surrounding host requires localized probe techniques. Atomic force microscopy, for example, can reveal a wealth of information about biogenic crystal growth processes (*8*, *9*). But how can we directly probe the interfaces between endogenous biopolymers and their inorganic hosts?

High-angle annular-dark-field scanning transmission electron microscopy (HAADF-STEM) is particularly well suited to the three-dimensional tomographic imaging of relatively thick crystalline specimens at resolutions of 1 nm (*10*, *11*). Unlike bright-field images, which can be complicated by electron-optical effects such as coherent diffraction, the HAADF signal is essentially incoherent (*12*). Thus, samples of different thicknesses (and therefore orientations) can be compared without elaborate corrections, allowing three-dimensional image reconstruction from a series of angular slices.

Li and co-workers used STEM, electron diffraction, and HAADF-STEM tomography to probe a model system in which single crystals of calcite are formed within a matrix of agarose, a gel-forming polysaccharide. Earlier work (*13*) had shown that agarose gels with high shear strength survive within the growing mineral phase, whereas weaker gels are excluded or broken apart. At length scales of hundreds of nanometers, the rigid agarose occurs as a random network of polymer aggregates that extend throughout the bulk of this single crystal.

The external morphology of these calcite-agarose composites is the ordinary rhombohedral structure with its six equivalent faces. However, the HAADF-STEM tomographs reveal that on the length scale of tens of nanometers, other high-energy facets occur at the endohedral interfaces between the polymer aggregates and the mineral phase. These calcite facets identified by Li *et al.* are rarely observed in the absence of polyanionic surfaces (*14*).

How are these high-energy interfaces formed? Although we ordinarily look for molecular recognition processes that control faceting, this composite system presents a different possibility: Once the polymers are trapped within the growing calcite structure, restoring forces within the polymer chains overcome the propensity to form only the most

Department of Chemistry, 213 CBC Building, Kansas State University, Manhattan, KS 66506–0401, USA. E-mail: mdholl@ksu.edu

stable interfaces. The rheology of the incorporated biopolymers, not molecular recognition, may very well be the critical factor in determining which internal facets are expressed in single crystals of bioinorganic composites.

Given the dire consequences of ocean acidification (*15*) and the enormous problem of CO_2 sequestration, the carbonates (e.g., $CaCO_3$) may be the most ecologically important minerals of the century. Approaches such as the ones described by Li *et al.* will be essential for a deeper understanding of $CaCO_3$ crystallization (and its dissolution) in biological systems.

References and Notes

1. B. L. Smith *et al.*, *Nature* **399**, 761 (1999).
2. H. Li, H. L. Xin, D. A. Muller, L. A. Estroff, *Science* **326**, 1244 (2009).
3. A. Berman *et al.*, *Science* **259**, 776 (1993).
4. J. Aizenberg, J. Hanson, T. F. Koetzle, S. Weiner, L. Addadi, *J. Am. Chem. Soc.* **119**, 881 (1997).
5. K. Henriksen, S. L. S. Stipp, J. R. Young, M. E. Marsh, *Am. Mineral.* **89**, 1709 (2004).
6. L. Addadi, S. Weiner, *Nature* **411**, 753 (2001).
7. B. Kahr, R. W. Gurney, *Chem. Rev.* **101**, 893 (2001).
8. T. Jung *et al.*, *Langmuir* **20**, 8587 (2004).
9. C. A. Orme *et al.*, *Nature* **411**, 775 (2001).
10. P. A. Midgley, R. E. Dunin-Borkowski, *Nat. Mater.* **8**, 271 (2009).
11. P. A. Midgley, M. Weyland, *Ultramicroscopy* **96**, 413 (2003).
12. J. M. Thomas *et al.*, *Angew. Chem. Int. Ed.* **43**, 6745 (2004).
13. H. Li, L. A. Estroff, *Adv. Mater.* **21**, 470 (2009).
14. J. Aizenberg, A. J. Black, G. M. Whitesides, *Nature* **398**, 495 (1999).
15. J. C. Orr *et al.*, *Nature* **437**, 681 (2005).

16. This work was supported by the NSF (grant CHE-0809845).

10.1126/science.1183122

CELL BIOLOGY

Nuclear Export of Small RNAs

Murray Stewart

The machinery that exports small RNAs from the cell nucleus is fine-tuned to recognize when the processing of each class of RNA has been completed.

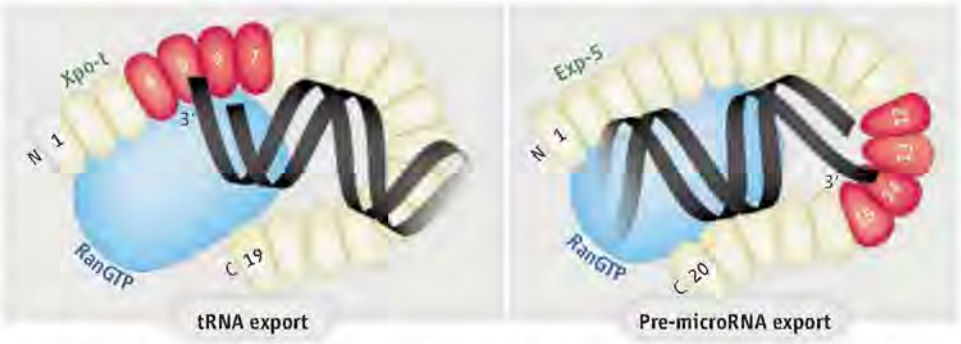
In eukaryotic organisms, RNAs are generated in the cell nucleus but many perform their functions in the cytoplasm and so need to be exported through pores in the nuclear envelope. Because many RNAs must be processed or modified in the nucleus, it is important that only the fully mature RNA is exported. Stringent molecular recognition is required to identify the class of RNA and ensure that its processing has been completed. On page 1275 of this issue, Okada *et al.* provide remarkable insight into this export process for pre-microRNAs (*1*), which—after further digestion in the cytoplasm—function to regulate specific mRNAs. Unexpectedly, pre-microRNAs are recognized for export in a quite different way to that recently reported by Cook *et al.* for transfer RNA (tRNA) (*2*).

Large mRNAs and a number of smaller RNAs—including tRNA and pre-microRNA—are transported to the cytoplasm through nuclear pores by nuclear export factors, or exportins, that are specific for each class of RNA (*3*). Both pre-microRNA and tRNA are exported by exportins of the β-karyopherin superfamily, in conjunction with the guanosine triphosphatase (GTPase) Ran. Pre-microRNA is exported primarily by exportin-5 (Exp-5) and tRNA by exportin-t (Xpo-t).

The export of pre-microRNA and tRNA employs the general Brownian ratchet model seen with other nuclear trafficking pathways (*4*). This model is based on the nucleotide state of Ran, which binds guanosine triphosphate (GTP) in the nucleus but guanosine diphosphate (GDP) in the cytoplasm. Exportins bind their cargoes together with GTP-bound Ran (RanGTP) in the nucleus. Weak interactions between nuclear pore proteins (nucleoporins) and the exportins mediate equilibration between the nucleus and cytoplasm though nuclear pores. When the export complex reaches the cytoplasm, it dissociates, releasing the cargo. The RanGTP is then hydrolyzed (providing the energy to power the process), and the Ran and exportins are recycled to the nucleus to participate in another export cycle.

Both pre-microRNA and tRNA are generated from larger precursor RNAs that undergo considerable nuclear processing. The exported forms are both based on stem loops with 3′ overhangs that are crucial for recognition by the export machinery. After transcription, capping, and polyadenylation, primary microRNAs are cleaved by the Drosha ribonuclease to generate pre-microRNA stem loops containing about 65 nucleotides that have a two-nucleotide 3′ overhang. tRNAs are also cleaved from a larger precursor that generates a three-nucleotide (CCA) 3′ overhang, located on the "acceptor helix" of the L-shaped tRNA molecules, where amino acids are attached in aminoacyl-tRNAs. Probably for quality control, export generally occurs only after this step has been completed. Because nuclear processing of both tRNA and microRNA needs to be completed before they are exported, an effective molecular recognition is required to distinguish between the mature molecules and their various precursor forms.

Exp-5 and Xpo-t are horseshoe-shaped molecules constructed from HEAT repeats (see the figure). HEAT repeats are based on two antiparallel α helices that generate molecules resembling helical springs (*4*, *5*). Like other karyopherins (*5*), these springlike molecules are intrinsically flexible, which is crucially important for both cargo binding and release (*2*). In the cargo-bound state, RanGTP



How tRNA and pre-microRNA bind to their exportins. The exportins are constructed from HEAT repeats (yellow and red) and adopt a roughly U-shaped conformation. An RNA double helix (black) is bound along most of the exportin's length, interacting mainly with positively charged residues on the exportin (for clarity, only the acceptor helix of tRNA is shown). RanGTP (blue) binds to HEAT repeats at either end of the exportin, helping to maintain the U-shaped conformation. In both exportins, the 3′ end is buried in a cavity formed by a group of HEAT repeats (red), yet the orientation of the tRNA helix (*2*) is opposite to that seen with pre-microRNA (*1*).

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. E-mail: ms@mrc-lmb.cam.ac.uk

interacts with regions near the N and C termini of Xpo-t and Exp-5 to clamp them around a double-stranded RNA helix (the tRNA acceptor helix or the pre-microRNA stem, respectively). This configuration facilitates interactions between positively charged residues on the inner surface of the exportin horseshoe and negatively charged RNA backbone phosphates. This binding is independent of the sequence of the RNA, enabling the exportins to bind a range of different RNAs.

Recognition of 3′ overhangs is crucial for the function of both Exp-5 and Xpo-t, but this is achieved in quite different ways, and remarkably, the RNA helix that is gripped by the karyopherin runs in opposite directions (see the figure). In Exp-5, the two-nucleotide 3′ overhang is recognized by a deep tunnel-like cavity formed by HEAT repeats 12 to 15 (*1*); in contrast, the tRNA 3′ CCA overhang is recognized by a groove formed by the inter-repeat sequences between HEAT repeats 4 and 7 of Xpo-t (*2*). In both pre-microRNA and tRNA, steric clashes involving the 5′ end seem to be important in recognizing the correct length of the 3′ overhang.

Both tRNAs and pre-microRNAs are recognized primarily by the 3′ overhang, which poses the question of why a separate export pathway has evolved for pre-microRNA. One possible explanation relates to the stability of these different classes of RNA. tRNAs have a compact, tightly folded structure and so are comparatively stable. By contrast, while they remain in the nucleus, pre-microRNAs are susceptible to digestion by ribonucleases that digest RNAs from either end. Many other RNAs are protected by modifications at their ends (such as adding a 5′ cap and 3′ polyA), but these are removed when pre-microRNAs are generated from primary microRNAs. Exp-5 grasps its cargoes as in a mitt, resulting in an intimate contact that may offer greater protection from digestion prior to export than would be afforded by binding to Xpo-t.

This protective feature may be important in a wider context because, in addition to pre-microRNAs, Exp-5 also exports other small RNAs, such as human Y1 RNA and adenovirus VA1 RNA. It seems likely that all these cargoes are recognized in similar ways, although the length of the 3′ overhang appears to influence the strength of each interaction (*1*).

Overall, the structures of both Exp-5 and Xpo-t show analogous recognition strategies to ensure that nuclear processing of their cargoes has been completed before they are exported. However, although burying the 3′ overhang is central to both, the precise way in which Exp-5 and Xpo-t achieve this differs in ways that facilitate protection of pre-microRNAs and accommodation of tRNA aminoacylation.

References and Notes

1. C. Okada *et al.*, *Science* **326**, 1275 (2009).
2. A. G. Cook *et al.*, *Nature* **461**, 60 (2009).
3. A. Köhler, E. Hurt, *Nat. Rev. Mol. Cell Biol.* **8**, 761 (2007).
4. M. Stewart, *Nat. Rev. Mol. Cell Biol.* **8**, 195 (2007).
5. E. Conti *et al.*, *Curr. Opin. Struct. Biol.* **16**, 237 (2006).
6. Supported in part by a Wellcome Trust Programme Grant.

10.1126/science.1183273

VIROLOGY

Moving Forward in HIV Vaccine Development

Clinical trial results of a vaccine regimen to prevent HIV infection show modest protection and suggest that changes are needed in the direction of future research.

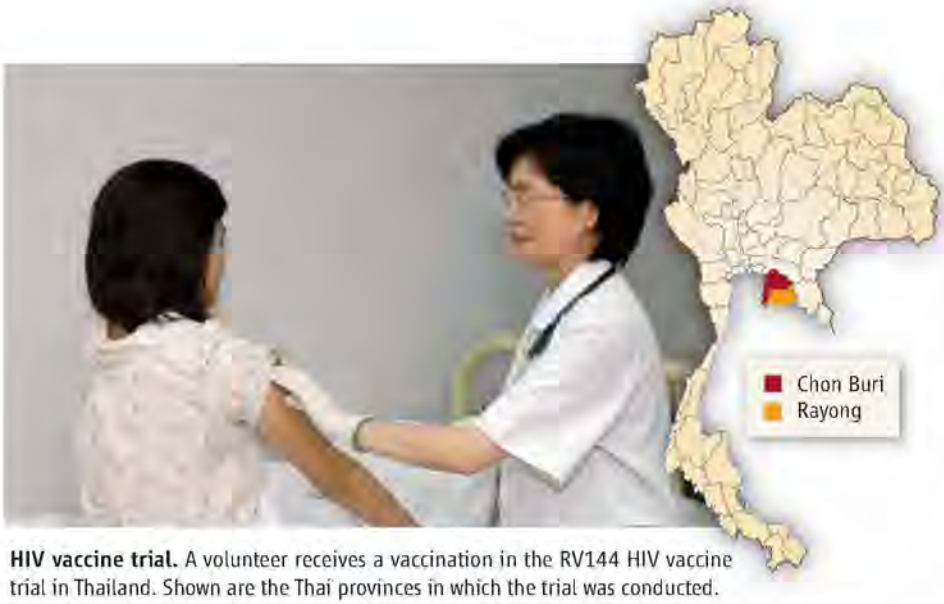
Norman L. Letvin

Early in the AIDS epidemic, efforts to develop a vaccine to prevent HIV infection were focused on two vaccine strategies. It was hoped that the HIV envelope glycoprotein (gp120) would generate an antibody response that would block the initiation of HIV infections, and that a recombinant canary pox construct expressing HIV genes would elicit cellular immune responses that would inhibit HIV replication. However, the nature of the immune responses elicited by each of these vaccine candidates in monkeys and human subjects proved disappointing (*1*–*3*). Despite these results, the U.S. Department of Defense (DOD) proceeded with plans to initiate the RV144 trial in Thailand to test the efficacy of a vaccine regimen that included both agents. At that time, the U.S. National Institutes of Health assumed responsibility for a major component of the DOD HIV/AIDS program, and provided a substantial proportion of the funding for the Thai trial.

Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, 02115, USA. E-mail: nletvin@bidmc.harvard.edu

The decision to commit resources to carry out the RV144 trial in the face of the limited immunogenicity of both vaccine prototypes created considerable concern within the HIV vaccine research community (*4*). Opponents of the trial argued that the resources being dedicated to the vaccine study would be better spent on basic research that might inform HIV vaccine development. As a consequence of discussions among HIV vaccine investigators, a viral load set point—the amount of virus present in the blood during chronic infection—was added as one of the primary endpoints (an outcome that can be measured objectively to determine whether the vaccination was beneficial) in the RV144 trial a year after its initiation. I shared the concerns of my colleagues about the wisdom of proceeding with this clinical trial, but my opposition was tempered by an appreciation for the commitment made by the DOD and the government of Thailand to test this vaccine.

It is unfortunate that there was a partial release of data before the formal announcement and publication of the trial findings, but we can now digest the complete results in the recently published report of the trial (*5*). The basic message is that the trial results show modest vaccine efficacy at 3 years after initiating vaccination, with a lowered rate of HIV infection by 31.2% ($P = 0.04$)—51 infections in the 8197 vaccine recipients and 74 infections in the 8198 placebo recipients. This evaluation was done using a modified Intent-to-Treat analysis, in which all uninfected participants enrolled in the trial were considered, including those who missed a shot or did not receive one at the appropriate time. The modest statistical significance of the finding of protection is lost, however, if a different strategy—a Per Protocol analysis—is employed for determining which study volunteers to analyze. In the Per Protocol analysis, only a subset of trial participants are considered—those who had followed the study design precisely and who also did not become infected during the 6-month period of vaccination. Both of these commonly employed analytic strategies were formally designated as the criteria for evaluating the study data before the trial was initiated, and therefore it is appropriate to evaluate the trial



HIV vaccine trial. A volunteer receives a vaccination in the RV144 HIV vaccine trial in Thailand. Shown are the Thai provinces in which the trial was conducted.

CREDIT: U.S. MILITARY HIV RESEARCH PROGRAM

results using both analytic approaches. The loss of the modest statistical significance of the findings using the Per Protocol analysis is likely a consequence of the loss of power in the study resulting from the smaller number of subjects that were analyzed.

Interestingly, in those participants who became infected after vaccination, no better control of HIV replication and no better protection against loss of the $CD4^+$ subset of T lymphocytes were observed when compared with the placebo recipients. Critics have pointed out that the absence of a vaccine effect on HIV replication in vaccinees who became infected contrasts with the findings in numerous HIV vaccine trials in nonhuman primates in which virus replication is partially contained in vaccinated animals that become infected.

The findings in the RV144 trial do not bear out the arguments of either the early supporters or opponents of the trial. The supporters argued at the outset that the study would provide an opportunity to test the utility of a "prime-boost" vaccine strategy that elicits both humoral and cellular immune responses (*6*). However, because the immunogens used in this trial did not generate antibodies that neutralize diverse HIV isolates or potent responses by $CD4^+$ or $CD8^+$ T cell subsets, and because the protection against HIV acquisition was very modest, the results of the trial cannot be viewed as either supporting or arguing against this prime-boost vaccine strategy. The trial opponents argued that the immunogenicity of the vaccine regimen was too marginal to warrant its testing in an efficacy trial (*4*). Yet, if the results of the Thai vaccine trial are viewed as hypothesis-generating rather than a step toward vaccine licensure, there are indeed some findings that could catalyze new efforts.

Although the number of subjects evaluated does not allow an assessment of the statistical significance of the finding, the vaccine may have been more effective immediately following administration, and this effect may have diminished rapidly as the weak vaccine-elicited immune responses waned. Immunogens that elicit more durable immune responses might therefore generate more durable protection.

Also of interest is the suggestion that the risk for HIV infection among the volunteers may have had a substantial impact on the results of the trial. Although the study was not sufficiently powered to assess vaccine efficacy in subpopulations of volunteers, it is provocative that the vaccine appeared 40% effective in the lower-risk population (limited number of low-risk sex partners and no needle exposure) and conferred only 3.7% protection in the higher-risk population (needle sharing or high-risk sexual activity). We will need to determine whether recombinant gp120 or recombinant canary pox used alone might confer protection against infection in a low-risk human population. A vaccine that elicits more potent responses may be needed to confer protection in high-risk populations.

However, it is also possible that the findings in the RV144 trial may not provide a useful lead for developing an effective immunization strategy for preventing HIV acquisition in high-risk populations. A vaccine that elicits qualitatively different immune responses may be needed to confer protection in high-risk individuals and may ultimately prove the best vaccination approach for a general population.

It will be important going forward to evaluate the immune responses elicited by recombinant protein and recombinant pox virus immunizations to explore potential immune correlates of protection in the RV144 trial. Because a large number of the immunized individuals were likely not exposed to HIV during the course of the vaccine trial, the most useful approach for exploring this problem will be to evaluate vaccinees who became infected and explore what immune responses were not elicited in them. These studies will be most informative if they are done using robust assays with wide dynamic ranges. The DOD has established advisory committees in the areas of B lymphocytes, T lymphocytes, genetics, and animal models to provide guidance on what studies should be done with the limited available blood specimens collected from the subjects in this study. There are very limited quantities of cells and no mucosal specimens available from the vaccine trial recipients (*7*). It will therefore be necessary to initiate new, focused immunogenicity studies in the near future to evaluate the cellular and humoral immune responses in systemic and mucosal compartments induced by the vaccine regimen administered in the RV144 trial. Based on our current understanding of the immune responses that are elicited by the immunogens used in this trial, early thinking is focusing on possible contributions to protection by antibodies that bind to but do not neutralize the virus and by $CD4^+$ T lymphocyte populations that augment vaccine-induced antibody responses to the HIV envelope glycoprotein (gp120).

The findings in the Thai trial suggest that different criteria should be considered in the future for determining what vaccine strategies warrant evaluation in advanced-phase clinical testing. Until now it was assumed that blocking HIV acquisition through vaccine-elicited immunity might not be achievable. Therefore, advanced-phase clinical studies have been initiated to determine whether vaccine-elicited immunity can contribute to decreasing the viral load after HIV infection has occurred. Because evidence from the Thai trial suggests that vaccination might result in protection against HIV acquisition, at least in low-risk subjects, clinical efficacy trials of vaccine candidates should in the future be initiated only if there is an expectation that protection from infection might be achieved. Therefore, HIV vaccine candidates being considered for advanced-phase human trials should be evaluated in nonhuman primates to determine whether they confer protection against HIV acquisition. This represents a change in the way nonhuman primate vaccine trials have often been done in the past. In those earlier studies, virus replication was monitored after high-dose, intravenous challenge. A more accurate prediction of vaccine efficacy will require the development of nonhuman primate models in which vaccine protection against the acquisition of HIV infection can be reliably evaluated. These models

would ideally be based on repeated, low-dose mucosal exposure of nonhuman primates to a virus. Useful studies can be done using pathogenic simian immunodeficiency viruses (SIVs) for these challenges. Because HIV envelope protein vaccines must be evaluated, it will also be important to develop consistently pathogenic chimeric simian-human immunodeficiency viruses—laboratory-constructed viruses that express the HIV envelope on an SIV "backbone" of structural and enzymatic proteins—that bind to CCR5, the human HIV co-receptor.

In retrospect, modifications in the design of the RV144 trial would have elucidated many of the findings with which we are now grappling. The study of larger numbers of individuals would have clarified the significance of the differences in HIV acquisition observed between the vaccine and placebo recipients. The delivery of the gp120 glycoprotein with a more potent adjuvant, such as one incorporating an agonist of Toll-like receptors (when activated, these receptors trigger immune cell responses), might have potentiated the protein component of vaccine. Similarly, the use of a less attenuated pox virus–based vector such as the NYVAC or MVA strains of vaccinia virus might have boosted immune responses. Also, it would have simplified the interpretation of the data if a second placebo arm in the trial had received the canary pox vector without an HIV gene insert, as this would allow a determination of the possible contribution of vector-induced innate immune responses to protection. Including such a second placebo arm in the trial could only be done if the U.S. Food and Drug Administration determined that the vaccine vector had a sufficiently benign safety profile.

While the pursuit of basic laboratory research will continue to be important in the development of an HIV vaccine, the results of the Thai trial underscore the extraordinary importance of also performing focused human clinical trials of vaccine strategies. Just as the nature of the recent failure of a recombinant adenovirus vaccine to confer protection against HIV infection in the STEP trial could not have clearly been predicted based on the preclinical experiments that had been carried out (*8*), the findings in the Thai trial were not expected based on preclinical studies and human immunogenicity data. We have learned to expect the unexpected in our efforts to generate an effective HIV vaccine.

References

1. J. Mascola *et al.*, *J. Infect. Dis.* **173**, 340 (1996).
2. N. M. Flynn *et al.*, *J. Infect. Dis.* **191**, 654 (2005).
3. N. D. Russell *et al.*, *J. Acquir. Immune Defic. Syndr.* **44**, 203 (2007).
4. D. R. Burton *et al.*, *Science* **303**, 316 (2004).
5. S. Rerks-Ngarm *et al.*, *N. Engl. J. Med.* (10.1056/NEJMoa0908492 (2009).
6. R. Belshe *et al.*, *Science* **305**, 177 (2004).
7. N. Michael, personal communication
8. S. Buchbinder *et al.*, *Lancet* **372**, 1881 (2008).

10.1126/science.183278

MICROBIOLOGY

Variety—the Splice of Life—in Microbial Communities

Splicing of viral fragments into bacterial and archaeal genomes may provide protection against viral predation.

Jillian F. Banfield[1] and Mark Young[2]

Bacteria and Archaea are commonly thought to exist as clonal populations because they reproduce asexually. However, genomic studies have shown that the evolutionary trajectories of most bacterial and archaeal populations are profoundly modified by exchange of genetic information, which creates natural populations that are inherently diverse. Much of this genomic variation confers no selective advantage or leads to evolutionary dead ends, some provides the raw material for selection that gives rise to new capabilities and lineages, and some may be critically involved in defense against mobile elements. Recently, it has become apparent that many Bacteria and Archaea splice into their genomes tiny fragments from the viruses and other mobile elements that harness them to replicate and survive. The uptake of these fragments can be so rapid that most bacterial and archaeal cells in a single population could be genomically distinct at the hypervariable locus where fragment insertion occurs. Variety that results from this diversification process could, in some cases, be critical to host population survival.

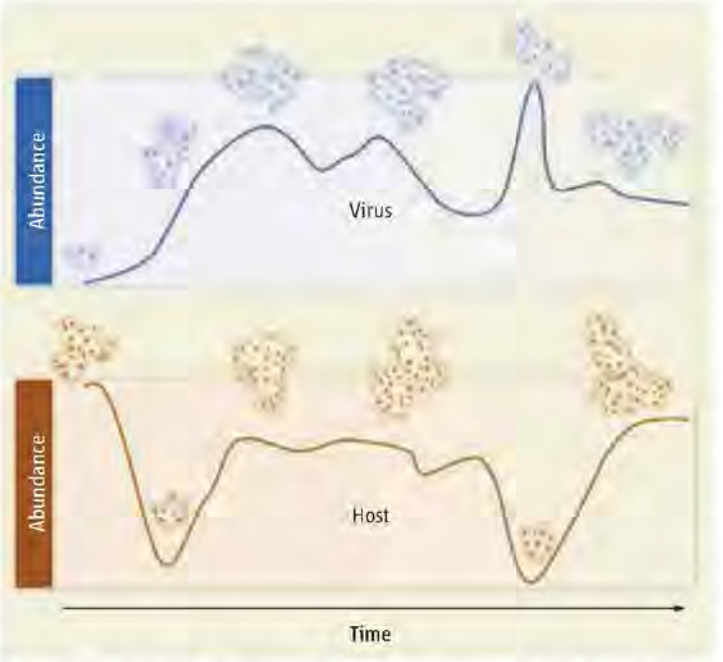
The small genome fragments—derived mostly from viruses of archaea and bacteria (phage), plasmids, and transposons—are inserted between repeat elements of regions called CRISPR loci (*1*–*3*). After the repeats and intervening spacers that comprise a CRISPR locus are transcribed and diced, the spacers serve as small RNAs (crRNAs) that can interfere with foreign nucleic acids because they share sequence identity with them (*4*). The crRNAs are flanked by partial repeats that function in conjunction with a subset of CRISPR-associated (Cas) proteins to target invading nucleic acid, either DNA or RNA, depending on the CRISPR/Cas system. Many Bacteria and Archaea possess one or more CRISPR/Cas system (*5*, *6*). Given that CRISPR/Cas systems can be transferred by mobile elements, we consider it likely that most lineages have had such systems over a large part of their history. Rapid progress toward understanding the machinery of these remarkable systems has been made, but appreciation of the impact of CRISPR/Cas in natural environments has lagged behind.

Metagenomic analyses of DNA fragments recovered from the environment provide snapshots of genomic variation within natural populations (*1*, *7*–*12*). This variation is evident through comparison of the sequences of coexisting individuals, and provides insight into the immune potential and spacer acquisition history of the population, as well as the countermeasures that viruses must engage in to maintain their virility. This same variation can be monitored in laboratory experiments in which microbes are challenged by viruses that can infect them.

Both approaches show that CRISPR loci play a major role in providing resistance to viruses (and other mobile elements), and that sequential, unidirectional addition of spacers coding new crRNAs to a CRISPR locus generates a "tape recording" of immune history (*1*, *2*). Viral populations respond to acquisition of a CRISPR spacer by a combination of mutation, homologous recombination, and deletion of the target sequence. Muta-

[1]Departments of Earth and Planetary Science and Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA. [2]Departments of Plant Sciences and Microbiology, Montana State University, Bozeman, MT 59715, USA. E-mail: jbanfield@berkeley.edu; myoung@montana.edu

tions that eliminate spacer efficacy can be either in the target region of the viral genome or in the flanking motif (*10*, *13*) that may be involved in sequence sampling, targeting, or both (*5*). Patterns of spacer organization in CRISPR loci in natural populations resemble patterns generated in laboratory experiments. Both experimental and metagenomic analyses indicate that rapid viral evolution to evade newly acquired crRNAs is countered by rapid acquisition of new spacer sequences by the host. This sets in motion simultaneous diversification of hosts and their mobile element predators, generating "cloudlike" diversity patterns in natural populations (see the figure). Over time, these clouds, representing ranges of infectivity and resistance, are modulated by interplay between the host and virus populations. These complex interactions are likely to affect overall microbial community structure and may be key to both host and virus population survival.



Dynamic equilibrium. In addition to well-known mechanisms of defense against viruses and other mobile elements, bacteria and archaea have an adaptive CRISPR/Cas-based immune system with stochasticity at its heart. Key to survival of both host and infective agent is population diversity, represented by "clouds." Over time, the diversity clouds morph (reflecting changing genomic composition), shrink (due to loss of diversity), and expand, and population size increases and decreases (lines). High levels of diversity in immunity and infectivity can be lost in bottlenecks due, for example, to the appearance of a particularly virulent virus (first dip in the host curve), but are regenerated rapidly. Overall, there is potential for a dynamic equilibrium between bacteria and archaea and their infective agents.

Acquisition of new spacers can lead to rapid diversification of the complement of crRNAs within the host population. If the viral population is very diverse and virus particles are numerous compared to the host cells (as expected), then many spacers may be unique in the host population, and the set of crRNAs (and hence the immune potential) of each individual could be different. However, it has been observed that all individuals can have essentially the same complement of older spacers, with a transition from shared to divergent sets of crRNAs near the tip of the locus where most recently added spacers reside (*1*). Such an observation suggests a recent strong selection event, followed by population diversification. Such a "sweep" could be caused by the appearance of an unusually virulent virus for which potentially only one cell in the population had immunity, or a locus transfer event from another population or from a plasmid. Where the host population has spacers with differing levels of efficacy, selective sweeps could result in multiple different crRNA complements, with dominance by the locus type with the most effective spacer or spacers. Although the data on hand are limited, initial indications are that some bacterial and archaeal population histories can be frequently—perhaps on the time scale of weeks, months, or years—punctuated by strong selection events. Interestingly, it has been proposed recently that CRISPR/Cas are fine-tuning for viral defense systems based on cell surface phage recognition (*14*). The widespread genomic distribution of CRISPR/Cas loci and their exceedingly rapid diversification in laboratory and perhaps natural environments compared to all other genomic change indicate that CRISPR/Cas are often, and perhaps usually, a primary line of defense against viral predation.

The host and virus diversification processes appear to be somewhat stochastic, and this may have important evolutionary consequences. Only after a virus particle penetrates a host cell will it be determined whether the host can suppress it via a CRISPR/Cas-based mechanism. The outcome of infection is thus a matter of probability. Further, if the spacer set of each cell results from a series of essentially random acquisition events, immune levels are probably independent of other host genome features. Thus, strong virus-driven selection events based on short spacer sequences just 23 to 47 nucleotides in length could sweep from the host population genomic variation that otherwise could lead to evolutionary or ecological advantage. It is unknown if restriction of selection events to geographically local subpopulations prevents loss of diversity. Virus-related bottlenecks could potentially have a significant impact on the physiological properties and fitness of the host, but it is unclear whether they can influence species evolutionary trajectories.

An intriguing feature of the CRISPR/Cas system is the large number of Cas protein families and the high levels of sequence variation within these families. This probably reflects rapid evolution driven by high selective pressure on the defense system, perhaps accelerated over periods when loci are encoded on genomes of mobile elements. The extensive phylogenetic distribution and wide diversity may also point to an ancient origin for the CRISPR/Cas system (*15*). For example, the CRISPR/Cas system may have arisen in plasmids, perhaps originally for plasmid-plasmid competition. Candidates for the ancestral protein machinery might be plasmid proteins involved in replication, DNA transfer, and integration. If evolved in such a mobile element pool, the systems appear to have been co-opted in modern organisms more for silencing of viruses than plasmids. Regardless of their origin and the extent to which CRISPR/Cas and virus co-evolution imparts population stability or leads to a history punctuated by bottlenecks, we anticipate that its functioning at the population level will motivate a shift in the paradigm for microbial community dynamics.

References and Notes

1. G. Tyson, J. F. Banfield, *Environ. Microbiol.* **10**, 200 (2008).
2. R. Barrangou *et al.*, *Science* **315**, 1709 (2007).
3. L. A. Marraffini, E. J. Sontheimer, *Science* **322**, 1843 (2008).
4. S. J. Brouns *et al.*, *Science* **321**, 960 (2008).
5. K. S. Makarova *et al.*, *Biol. Direct* **1**, 7 (2006).
6. D. H. Haft *et al.*, *PLOS Comput. Biol.* **1**, e60 (2005).
7. A. F. Andersson, J. F. Banfield, *Science* **320**, 1047 (2008).
8. L. J. Wilhelm *et al.*, *Biol. Direct* **2**, 27 (2007).
9. G. J. Dick *et al.*, *Genome Biol.* **10**, R85 (2009).
10. J. F. Heidelberg *et al.*, *PLoS One* **4**, e4169 (2009).
11. A. Djikeng *et al.*, *PLoS One* **4**, e7264 (2009).
12. R. Danovaro *et al.*, *Nature* **454**, 1084 (2008).
13. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
14. F. Rodriguez-Valera *et al.*, *Nat. Rev. Microbiol.* **7**, 828 (2009).
15. K. S. Makarova *et al.*, *Biol. Direct* **4**, 29 (2009).
16. We thank R. Barrangou for his input to this perspective. Funding support from the Department of Energy Genomic Science Program and NASA Astrobiology Institute, and meeting support from the Army Research Office is acknowledged.

10.1126/science.1181501

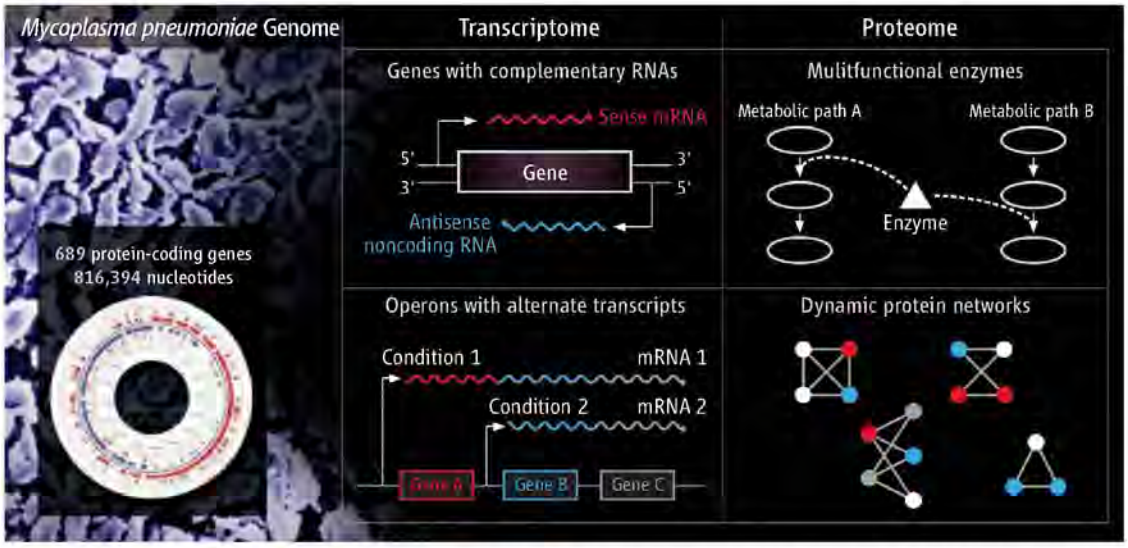
SYSTEMS BIOLOGY

Excavating the Functional Landscape of Bacterial Cells

Functional analyses of a reduced bacterial genome suggest levels of complexity and control previously assumed to be restricted to eukaryotes.

Howard Ochman and Rahul Raghavan

There is a poster that charts the metabolic pathways of the model bacterium *Escherichia coli*, and our laboratory uses it as a reference to convert each new bacterial genome sequence into an atlas of encoded functions. This is a particularly satisfying endeavor and has been best applied to those host-dependent bacteria whose highly reduced genomes contain a subset of the genes in *E. coli*: If a particular gene or pathway is eroded or absent, it is assumed that the bacterium is deficient in that activity and no longer has a need for the trait in its current circumstances. However, this linear mapping of genes to function rarely considers how a cell actually accomplishes the processes it has retained. Genes are viewed simply as performing a specified function despite the many internal and external factors that might affect their implementation. Three papers in this issue—by Güell *et al.* on page 1268 (*1*), Yus *et al.* on page 1263 (*2*), and Kühner *et al.* on page 1235 (*3*)—report features of transcriptional control and protein organization that are much more subtle and intricate than were previously considered possible in bacteria, and in many ways, appear similar to mechanisms in eukaryotes.



Compact but complex. Analyses of the RNAs and proteins of the bacterium *M. pneumoniae* raise questions about the complexity of regulatory mechanisms in organisms with small genomes.

Using *Mycoplasma pneumoniae*, one of the causative agents of what is sometimes called "walking pneumonia," the authors of all three studies performed comprehensive and detailed investigations on the organism's small genome that have greatly increased our understanding of bacterial cellular systems (see the figure). Although numerous technological developments assisted these analyses, much of their beauty comes from the organism itself. Unlike the >4000-kb genomes of the most popular bacterial model systems *E. coli* (*4*) and *Bacillus subtilis* (*5*), *M. pneumoniae* has a more tractable genome size of only 816 kb, encoding proportionately fewer genes and regulatory elements. But unlike other bacteria with diminutive genomes, the smallest of which are host-dependent symbionts of insects (*6*), *M. pneumoniae* can be cultivated in vitro, allowing manipulation and detailed scrutiny of external parameters on its physiology.

Güell *et al.* analyzed the expression patterns of the complete set of RNAs encoded in the *M. pneumoniae* genome, revealing a complex transcriptional landscape. A large fraction of protein-coding genes have RNAs encoded on the complementary strand of genomic DNA, suggesting a double-stranded RNA–based regulation of gene expression akin to a process observed in eukaryotes (*7*). Genome-wide profiling of transcripts produced under various growth conditions revealed increased or decreased expression of contiguous genes (depending on the context), progressively descending levels of transcripts within operons (an operon includes several genes that are controlled by common regulatory elements), and the production of alternative transcripts from a single locus, indicating that bacterial operons are much more dynamic and versatile than previously thought. The authors have begun to decode these complicated regulatory signatures by assembling groups of coexpressed genes and searching for sequence motifs common to each group. Although this approach identified some novel regulatory elements, it seems likely that many remain unrecognized. Together, these findings suggest the presence of a highly structured, multifaceted regulatory machinery, which is unexpected because bacteria with small genomes contain relatively few transcription factors (*8*, *9*).

Yus *et al.* integrated biochemical, structural, and computational information to reconstruct the metabolic network of *M. pneumoniae*, and then used this metabolic map to develop a culture medium containing the minimal requirements necessary for cultivating the bacterium. The metabolic pathways in *M. pneumoniae*, compared to those of bacteria with larger genomes, contain few branches or redundancies. However, its metabolic potential is augmented by the presence of numerous enzymes that perform multiple functions. And as observed in the RNA analyses of Güell *et al.*, the reduction in the metabolic network of *M. pneumoniae* is accompanied by intricate changes in the pattern of regulation, further suggesting a complex system that involves multifunctional enzymes, post-translational modifications of proteins, and chemical messengers.

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85718, USA. E-mail: hochman@email.arizona.edu

The genome-scale analysis of protein complexes in *M. pneumoniae* by Kühner *et al.* greatly expands our knowledge of protein-protein interactions within bacterial cells. Proteins often interact with one another to form functional complexes, and similar to the situation with eukaryotes (*10*), more than 90% of soluble proteins in *M. pneumoniae* serve as components of protein complexes. Surprisingly, the protein-interaction networks correlate poorly with genome organization and gene expression patterns—gene adjacency and coexpression were not good predictors of physically interacting proteins—again suggesting the presence of regulatory mechanisms not apparent from the compact genome structure.

How did these remarkable layers of gene regulation and the highly promiscuous behavior of proteins in *M. pneumoniae* arise? At first glance, these features may seem to be fine-tuned adaptations to the organism's current life-style, but this is not compatible with evidence for the reduced efficacy of selection that operates on the genomes of host-dependent bacteria. Bacteria that chronically associate with eukaryotic hosts undergo bottlenecks at the time of transmission, and such reductions in long-term effective population size result in a relaxation of selection genome-wide. This instigates the accumulation and fixation of deleterious mutations in seemingly beneficial genes due to genetic drift, as well as in those genes rendered superfluous in the nutrient-rich host environment (*11*). In both cases, disrupted genes are eliminated by the pervasive mutational bias favoring deletions that is present in all bacteria, thereby reducing genome size (*12*).

As genes are lost, their roles are fulfilled by the remaining genes, much like the members of a downsized office-staff who perform tasks that previously were carried out by former co-workers. As evidence of this process, the smallest cellular genome, currently 144 kb for the insect symbiont *Hodgkinia cicadicola* (*6*), encodes only 15 transfer RNAs (tRNAs) to specify the 20 amino acids required to synthesize proteins. It is difficult to see how this could be adaptive, or even possible, but presumably several tRNAs must be assuming multiple roles.

The reduced genome of *M. pneumoniae* belies an underlying eukaryote-like cellular organization replete with intricate regulatory networks and innovative pathways, revealing that there is no such a thing as a "simple" bacterium. The compound roles of individual genes and the need for additional regulatory mechanisms both may be hallmarks of reduced bacterial genomes, and the extraordinary information now available for *M. pneumoniae* sets a new standard for understanding systems-level questions about bacterial physiology and evolution.

References

1. M. Güell *et al.*, *Science* **326**, 1268 (2009).
2. E. Yus *et al.*, *Science* **326**, 1263 (2009).
3. S. Kühner *et al.*, *Science* **326**, 1235 (2009).
4. F. R. Blattner *et al.*, *Science* **277**, 1453 (1997).
5. F. Kunst, *Nature* **390**, 249 (1997).
6. J. P. McCutcheon, B. R. McDonald, N. A. Moran, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15394 (2009).
7. S. Katayama *et al.*, *Science* **309**, 1564 (2005).
8. R. Himmelreich *et al.*, *Nucleic Acids Res.* **24**, 4420 (1996).
9. S. Shigenobu *et al.*, *Nature* **407**, 81 (2000).
10. A. C. Gavin *et al.*, *Nature* **440**, 631 (2006).
11. C.-H. Kuo, N. A. Moran, H. Ochman, *Genome Res.* **19**, 1450 (2009).
12. C.-H. Kuo, H. Ochman, *Genome Biol. Evol* **1**, 145 (2009).

10.1126/science.1183757

COMPUTER SCIENCE

What Can Virtual Worlds and Games Do for National Security?

Virtual environments based on behavorial models allow analysts to explore different outcomes of proposed actions in military conflicts.

V. S. Subrahmanian and John Dickerson

Military planners have long used war games to plan for future conflicts. Beginning in the 1950s, defense analysts began to develop computer-based models to predict the outcomes of military battles that incorporated elements of game theory. Such models were often restricted to two opposing forces, and often had a strict win-lose resolution. Today, defense analysts face situations that are more complex, not only in that conflicts may involve several opposing groups within a region, but also in that military actions are only part of an array of options available in trying to foster stable, peaceful conditions. For example, in the current conflict in Afghanistan, analysts must try to estimate how particular actions by their forces—building schools, burning drug crops, or performing massive security sweeps—will affect interactions between the many diverse ethnic groups in the region. We discuss one approach to addressing this prediction problem in which possible outcomes are explored through computer-based virtual-world environments.

War games (*1*) are used to play out certain scenarios that an expert has designed. Partial information games (*2*) allow machine-derived models, such as stochastic opponent modeling agents (SOMA) (*3*), to guide the actions of players in the game based on stochastic decision rules and in the presence of partial information about the other players' situation. U.S. forces might use SOMA models to understand that a particular group might respond in one of a million ways, together with a probability distribution over those million responses.

Virtual worlds provide a software environment within which players can virtually "see" these scenarios play out in front of them, understand the probabilities of the scenarios, understand what types of things another player might do, and explore "what ifs" with respect to opponents. Decision-makers gain simulated "experience" to guide real-world decisions by enumerating different ways that such complex interactions might play out over an extended period of time.

Underlying the output of a virtual world is a game tree, that is, a tree in which each node represents a "belief state." A belief state is a probability distribution over states, where a state is a set of conditions that are true in a particular node. A state may include not only what a group might be thinking, but also what its offensive capabilities are, where it has forces, or whether opponents are actively engaging them.

Each outcome of a player's actions drops the game a level in a tree. For example, the U.S. might play, with a group playing at the

Department of Computer Science and Institute for Advanced Computer Studies (UMIACS), University of Maryland, College Park, MD 20742, USA. E-mail: vs@cs.umd.edu; johnd@umiacs.umd.edu

A virtual-world terrain representation. In this example, an Afghan cultural island representing a town might be populated with realistic looking "avatars" representing Afghan men, women, and children whose actions are in accordance with behavioral models. If the town is 65% Pashtun, 30% Hazara, and 2% each of Tajik and Uzbek, with a smattering of other groups represented, then the avatars should follow this distribution as well. The accompanying video describes a set of different outcomes of analyst actions within this representation, which is an example where artificial groups are used.

next level, a second group playing at the level after that, and so forth until it is the U.S.'s turn to play again. A specific U.S. move will transition the belief state from one node to an outcome (or "child") node, but the outcomes cannot be known with certainty. A game tree can easily exceed $10^{1,000,000}$ nodes in size, even to model interactions over a short time frame involving a small number of groups. In the Afghanistan conflict (*3*), for example, different rates of improvised explosive devices attacks in one town represent several branches coming from one node in the game tree. New approaches that limit the number of options use a mix of computational behavioral models (*4*) and forecasting techniques (*4*, *5*) as well as game-theoretical algorithms (*6*). This hybrid analysis model leverages analysts' knowledge with a mix of gaming and virtual-world technology.

A number of virtual-world environments such as *Second Life*, *World of Warcraft*, and others have emerged as online phenomena. However, they do not provide direct support for game theory because neither platform includes real-world models of terrorist groups or sociocultural groups that their forces may encounter. In addition, these platforms do not provide explicit programming for game-theoretic reasoning, although characters can be inserted who perform according to game-theoretic reasoning. For instance, *Second Life* and *Olive* provide an environment where users can adopt a virtual persona that engages in various virtual activities, such as building a house, adopting a profession, and building friendships, but do not provide the interaction environment between groups. *World of Warcraft* provides an environment within which users engage in combat on a virtual battlefield that continuously changes conditions, but it does not consider why groups are fighting or how one group may reason about other groups.

Cultural islands developed at the University of Maryland (*7*), and shown in the figure, provide a virtual-world representation of a real-world environment or terrain, populated with characters from that part of the world who behave in accordance with a behavioral model. For example, the CONVEX (*8*) forecasting engine has been validated on historical data for about 100 groups and found to be more than 90% accurate. The CAPE forecasting engine (*4*), which focuses only on predicting behavioral changes, was validated on historical data for seven groups and had about 69% accuracy in predicting behavioral changes.

U.S. defense analysts can use such virtual worlds to interact with models of the behaviors of these groups and understand how certain actions they might take will affect the short-term and long-term behaviors of these groups. At any given point in time, the game has a "state" describing, for instance, the situation in a town. When U.S. forces or a local government takes actions such as opposing a local leader, that state is altered. A group may react in one of several ways in accordance with a probability distribution.

Although behavioral models can be rerun several times to check for robustness, the intent of the virtual worlds is to allow defense analysts to bring their own expertise into the game to explore how the U.S. or a particular group might act and look for favorable outcomes. Analysts can also make these virtual worlds more efficient by using their intuition to prune large parts of the enormous game tree. For instance, a behavioral model of the Pashtuns in a virtual world may display only the 20 most probable reactions to the analyst, which reduces the number of options to be searched by 98%. Thus, a human analyst playing the role of Hezb-i-Islami may examine the 20 most probable things this group might do and then choose just one of these options.

Defense analysts can understand the repercussions of their proposed recommendations for policy options or military actions by interacting with a virtual-world environment such as CIG (computational intelligence and games) (*7*), which already comes equipped with connections to behavioral models, or other sophisticated virtual-world environments, such as *Second Life* or *Olive* (*9*), both of which may be augmented with appropriate behavioral models. They can propose a policy option and walk skeptical commanders through a virtual world where the commander can literally "see" how things might play out. This process gives the commander a view of the most likely strengths and weaknesses of any particular course of action.

References

1. M. L. Herman, M. D. Frost, *Wargaming for Leaders: Strategic Decision Making from the Battlefield to the Boardroom* (McGraw Hill, New York, 2008).
2. R. J. Aumann, *Repeated Games with Incomplete Information* (MIT Press, Cambridge, MA, 1995).
3. S. Khuller *et al.*, *Ann. Math. Artif. Intell.* **51**, 295 (2007).
4. A. Sliva, V. S. Subrahmanian, V. Martinez, G. Simari, in *Mathematical Methods in Counterterrorism*, N. Memon, Ed. (Springer-Verlag, New York, 2009), pp. 253–269.
5. P. A. Schrodt, D. J. Gerner, *Am. Polit. Sci. Rev.* **94**, 803 (2000).
6. A. Kott, W. M. McEneaney, *Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind* (Chapman and Hall, London, 2006).
7. J. Dickerson, M. V. Martinez, D. Reforgiato, V. S. Subrahmanian, in *Proceedings of the Second International Conference on Computational Cultural Dynamics*, V. S. Subrahmanian, A. Kruglanski, Eds. (Association for the Advancement of Artificial Intelligence, Menlo Park, CA, 2008), pp. 26–31.
8. V. Martinez, G. I. Simari, A. Sliva, V. S. Subrahmanian, *IEEE Intell. Syst.* **23**, 51 (2008).
9. M. R. Macedonia, *IEEE Computer* **40**, 99 (2007).

10.1126/science.1182660

INTERNATIONAL

AAAS, Europe Build Engagement with S&T Policy, Cooperation Plans

ISPRA, Italy—The landscape is familiar to science policy experts on both sides of the Atlantic: To address serious global problems and to take advantage of important future discoveries, they will have to work effectively with elected officials and members of the public who have complex and sometimes conflicting values and interests.

Ambitious plans. Roland Schenkel (left), director-general of the European Commission's Joint Research Centre, and AAAS CEO Alan I. Leshner signed a 3-year agreement to pursue cooperative efforts on a range of S&T issues.

Nearly two dozen S&T policy leaders and stakeholders from Europe and the United States gathered here recently to explore how they can navigate that landscape to make their policy advice more effective. During a day of candid dialogue organized by the European Commission's Joint Research Centre (JRC) and AAAS, participants shared concerns and strategies for a time of global economic and policy challenges involving science and technology.

Public values and other social and economic considerations are as important as science in policy-making, the participants said. The key will be to build trust through integrity, transparency, and deeper engagement with policy-makers and the public. And on familiar issues such as climate change or issues arising from newer fields like neuroscience and synthetic biology, participants argued it is crucial for scientists to serve as honest brokers in assessing potential benefits and risks.

"What a unique setup," said JRC Director-General Roland Schenkel, who presided over the meeting with AAAS CEO Alan I. Leshner. "Twenty-two brilliant and experienced leaders coming from the scientific community, policy-makers, industry, and NGOs from both sides of the Atlantic, sitting together in a small room for a day, engaging in a lively, open debate without any taboos, and coming up with recommendations on the do's and don't's in providing science-based advice to policy-makers. It was a truly remarkable experience."

The high-level dialogue at the JRC campus north of Milan was emblematic of promising new efforts by the European Commission (EC) and AAAS to enhance the European-U.S. science cooperation that has endured for centuries.

This fall, the EC's delegation in the United States, the embassy of Sweden, and AAAS have organized a series of discussions in Washington, D.C., on adaptation to climate change.

For the longer term, AAAS will undertake a pair of 3-year EC projects. One, dubbed BILAT-USA and led by the Austrian Research Promotion Agency, will address key issues of trans-Atlantic importance in workshops and symposia and encourage American participation in cooperative research. The second, Link2US, will be led by AAAS with a goal of raising awareness in the European research community about cooperative opportunities in the United States.

The 27 October "Trans-Atlantic Science for Policy Workshop" in Ispra, where the EC's first research site opened 50 years ago, was the first of a 2-day AAAS-JRC engagement expected to yield long-lasting impact.

The discussion leaders were Patrick Cunningham, chief scientific adviser to the Irish government; Robert Madelin, EC director-general for health and consumer protection; John Vassallo, Microsoft vice president for European affairs; and Kerstin Niblaeus, senior adviser to Sweden's minister of environment. Among other participants were Marja Makarow, chief executive of the European Science Foundation; David Goldston, director of government affairs for the Natural Resources Defense Council; Vittorio Prodi, a member of the European Parliament; and Eugene B. Skolnikoff, professor emeritus at MIT.

The next day, in Brussels, Leshner delivered the inaugural lecture in an annual series planned by the JRC before an audience of more than 350 people, including top EC officials, science and technology leaders, business executives, and students.

In introducing Leshner, Janez Potočnik, European commissioner for science and research, envisioned a relationship based on "healthy competition" and strategic cooperation.

"Partnership is crucial if we are to succeed in tackling the undeniable, shared problems of climate change and energy, or food and security," Potočnik said. "We can—and we must—build a strong and long-standing relationship as international partners."

Leshner struck similar themes. Even as the United States and Europe join to solve grand challenges, he said, they must build global support for consistent standards on science education, research, and ethics. And, he added, they should share insights on building constructive engagement with the public.

"High-quality science now is going on all over the world," said Leshner, the executive publisher of *Science*. That underscores the "tremendous need and opportunity for us—the European Commission and the United States—to collaborate on integrating the global scientific community so that we can all take advantage of the benefits."

After the address, Leshner and Schenkel signed a 3-year agreement under which AAAS and the JRC will pursue cooperative efforts on a range of initiatives, with an initial focus on nuclear security issues.

COMMUNICATION

AAAS, Kavli Name Science Journalism Award Winners

A radio broadcast on probability told through a tale about a drifting balloon, a newspaper series on the impact of a devastating genetic disease on a family in rural Montana, and a group of gracefully written stories about genetics and evolution are among the winners of the 2009 AAAS Kavli Science Journalism Awards.

Large Newspaper—(Circulation >100,000): Carl Zimmer, *The New York Times,* for "Now: The Rest of the Genome," "10 Genes, Furiously Evolving," and "Blink Twice If You Like Me," 11 November 2008, 5 May 2009, and 30 June 2009.

Small Newspaper—(Circulation <100,000): Amie Thompson, *Great Falls (Montana) Tribune,* for "Lethal Legacy" (series), 21 to 23 June 2009.

Magazine—Gary Wolf, *Wired,* for "Barcode of Life," October 2008.

Television—(Spot News/Feature Reporting, ≤ 20 minutes): Julia Cort, NOVA scienceNOW, for "Diamond Factory," 30 June 2009.

Television—(In-Depth Reporting, >20 minutes): Doug Hamilton, WGBH/NOVA, for "The Last Extinction," 31 March 2009.

Radio—Jad Abumrad, Soren Wheeler, Robert Krulwich, WNYC Radiolab, for "A Very Lucky Wind," 15 June 2009.

Online—Lisa Friedman, ClimateWire, for "Bangladesh: Where the Climate Exodus Begins" (series), March 2009.

Children's Science News—Douglas Fox, *Science News for Kids,* for "Where Rivers Run Uphill," 23 July 2008.

The 2009 awards are the first given under a new endowment by The Kavli Foundation, which has ensured the future of a program that was established in 1945. The endowment has allowed expansion of the television category to include two awards for the first time.

Independent panels of science journalists select the winners of the awards. The winners for each category will receive $3000 and a plaque at the 2010 AAAS Annual Meeting in San Diego in February. —*Earl Lane*

ELECTIONS

Additional Candidates for AAAS Annual Election

The following candidates have been added to the ballot for the 2009 election of AAAS officers. Members registered in more than one section will receive ballots for elections for each section they are enrolled in. For a list of other candidates, please see AAAS News & Notes in the 25 September 2009 issue of *Science.*

In response to member requests, AAAS will be offering the opportunity to vote either by mail or via a Web balloting system for this year's election. The online option will provide AAAS with the opportunity for better communication during the election and make it more convenient for members to cast ballots. Detailed instructions for using the online option will be provided to all members.

Atmospheric and Hydrospheric Sciences
Chair Elect: Otis Brown, Univ. of Miami; Walter F. Dabberdt, Vaisala
Member-at-Large of the Section Committee: Mary Anne Carroll, Univ. of Michigan, Ann Arbor; William D. Collins, Lawrence Berkeley National Laboratory
Electorate Nominating Committee: James H. Butler, NOAA Earth System Research Laboratory; John W. Farrington, Woods Hole Oceanographic Institution; Ray F. Weiss, Scripps Institution of Oceanography; Russell R. Dickerson, Univ. of Maryland, College Park

Chemistry
Chair Elect: Cynthia J. Burrows, Univ. of Utah; Vicki Grassian, Univ. of Iowa
Member-at-Large of the Section Committee: Scott Rychnovsky, Univ. of California, Irvine; Ron Raines, Univ. of Wisconsin-Madison
Electorate Nominating Committee: Klavs Jensen, Massachusetts Institute of Technology; Viresh Rawal, Univ. of Chicago; Timothy P. Lodge, Univ. of Minnesota
Council Delegate: Jeanne Robinson, Los Alamos National Laboratory; Al Sattelberger, Argonne National Laboratory; John Lowe, JL3Pharma LLC; Clark Landis, Univ. of Wisconsin-Madison; Jack Norton, Columbia Univ.; John Hartwig, Univ. of Illinois, Urbana-Champaign

Dentistry and Oral Health Sciences
Chair Elect: Carolyn W. Gibson, Univ. of Pennsylvania School of Dental Medicine; Gary Armitage, Univ. of California, San Francisco
Member-at-Large of the Section Committee: Kenneth Yamada, National Institutes of Health; Richard Lamont, Univ. of Florida
Electorate Nominating Committee: Dennis Mangan, Univ. of Southern California; Anne George, Northwestern Univ.; Mina Mina, Univ. of Connecticut Health Center; Mark Lingen, Univ. of Chicago

Industrial Science and Technology
Chair Elect: Nominees to be announced
Member-at-Large of the Section Committee: Nominees to be announced
Electorate Nominating Committee: Nominees to be announced

Information, Computing, and Communication
Chair Elect: Vinton Cerf, Google; Lewis Branscomb, Harvard Univ.
Member-at-Large of the Section Committee: Maja J. Mataric, Univ. of Southern California; Walter Warnick, Office of Scientific & Technical Information, U.S. Dept. of Energy
Electorate Nominating Committee: John Peha, Carnegie Mellon Univ.; Jon Eisenberg, The National Academies; Tom Pyke, U.S. Dept. of Energy; Maureen Kelly, Consultant

Linguistics and Language Science
Electorate Nominating Committee: Kirk Hazen, West Virginia Univ.; Jean-Pierre Koenig, SUNY Buffalo; Don Ringe, Univ. of Pennsylvania; Joe Salmons, Univ. of Wisconsin-Madison

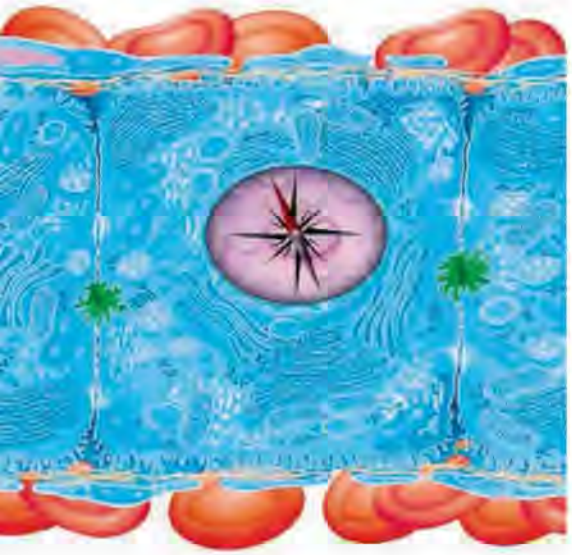
Societal Impacts of Science and Engineering
Chair Elect: Rachel Levinson, Arizona State Univ.; Jeremy Sugarman, John Hopkins Bloomberg School of Public Health
Member-at-Large of the Section Committee: Susanna Priest, Univ. of Nevada, Las Vegas; Howard Gobstein, Association of Public and Land-Grant Universities
Electorate Nominating Committee: Richard Johnson, Arnold & Porter LLP; Donna Nelson, Oklahoma Univ.; Anne Fitzpatrick, Dept. of Energy

2010 ELECTION

A Call For Nominations

AAAS members may suggest nominees (including themselves) for president-elect and the Board of Directors for election in the fall of 2010. For a list of this year's candidates, see AAAS News & Notes in the 25 September 2009 issue of *Science;* for a list of current Board members, go to www.aaas.org/aboutaaas/organization/board.shtml.

Please send the suggested nominee's curriculum vitae no later than 15 January 2010 to Gretchen Seiler, AAAS Executive Office, 1200 New York Avenue, N.W., Washington, DC, 20005. Suggested nominees will be considered by the AAAS Committee on Nominations at its winter meeting.



INTRODUCTION

Location, Location, Location

WHEN BUYING A HOUSE, ONLY THREE THINGS MATTER: LOCATION, LOCATION, location. In cell biology, a similar adage can be applied to the regulation of cellular and organismal physiology. The location of a cell within an organism and the location within the cell of its constituent parts will affect all it does, including the functions it is capable of performing, its signaling partners, and whether and how it grows and divides. Even in single-celled bacteria, spatial organization regulates cell division and other key developmental processes. In this special issue of *Science* we address a variety of topics that contribute to our understanding of spatial cell biology.

Chang (p. 1206) describes how position within the body affects a cell's differentiation and functional characteristics and how cells use specific gene expression programs to encode location. Pollard and Cooper (p. 1208) describe how a cell's internal actin cytoskeleton affects nearly all aspects of its biology, defining the cell's shape, controlling its movements, and contributing to cell motility and division. Holt and Bullock (p. 1212) go on to remind us that animal cells also regulate the intracellular localization of messenger RNAs, leading to localized translation of their encoded proteins. In metazoans, most cell types are housed within an extracellular matrix. Hynes (p. 1216) describes how the extracellular matrix and its constituent proteins do not simply act as passive supports but also actively influence their resident cells' physiology, integrating complex signals in space and time. Scott and Pawson (p. 1220) expand on this theme of regulation of signaling in space and time by examining the physical responses of intracellular signaling proteins as they transduce extracellular cues into intracellular effects. Finally, Shapiro, McAdams, and Losick (p. 1225) bring our attention to bacteria that also need to regulate their cellular anatomy by controlling the intracellular location of individual proteins and protein complexes during growth and division and in response to stress or other external cues.

Spatial cell biology affects all aspects of physiology in health and disease. For example, during limb development, the position of cells along a developmental axis regulates their differentiation to generate the appropriate appendage in the appropriate location. And as part of cancer metastasis, cancer cells develop the ability to escape their usual physical niche and set up residence elsewhere within the body, with often devastating consequences. In Victorian times, "knowing your place" was essential in the establishment and maintenance of social order. Within our bodies, this notion, outmoded for human societies, is essential for our physiological integrity.

– STELLA HURTLEY

Spatial Cell Biology

CONTENTS

PERSPECTIVE

Anatomic Demarcation of Cells: Genes to Patterns

Howard Y. Chang

An organizing principle of the diverse cell types in multicellular organisms is their anatomic location. In turn, anatomic location is patterned by the positional identities of cells along developmental axes. Recent progress in functional genomics and chromatin biology illustrates how cells use specific gene expression programs to encode location. Dynamic chromatin states of key genes, notably the *Hox* loci, serve as the internal representation in cells of their positional identity within the animal.

The same genetic blueprint gives rise to thousands of cell types that make up the human body, and these different cell types must be correctly arranged in spatial patterns to make functioning tissues and organs. In continuously regenerating tissues, the genome faces an additional challenge of ensuring the faithful transmission of information throughout a lifetime—in humans, over decades. An important organizing principle, based on the anatomic location of cells in the animal, guides many cell fate choices and is set in place during embryonic development (*1*). Thus, the key question is how cells in multicellular organisms "know" where they are located in the body. Recent studies have provided insight into how mammalian cells represent, remember, and act upon their positional identities.

Gene Expression and Positional Identity

Cellular identities are organized on the basis of position along developmental axes, for instance along the anterior-posterior (head to tail) axis or proximal-distal axis (close or far away from the trunk along limbs or other appendages) (Fig. 1A). Homeotic genes (*Hox* genes, encoding transcription factors) are expressed in a nested pattern along these developmental axes. *Hox* genes are master regulators of body morphogenesis in development and evolution, and mutations in *Hox* genes transform one body segment into another (*2*, *3*). But fundamental questions still remain. If a finger is defined by a combination of *Hox* genes that indicates anterior and distal positional identity, do those *Hox* genes together turn on one or more unique "finger" genes, or alternatively, does the combination of anterior and distal *Hox* target genes together shape a finger?

This conundrum started to be resolved with the advent of full-genome sequences and the technology to interrogate gene expression patterns genome-wide. Comparison of global gene expression patterns of cell types thought to be homogenous, but derived from different anatomic sites—including fibroblasts, endothelial cells, smooth muscle cells, fat cells, and bone—showed that each of these cell types retained extensive and stereotypic differences in their gene expression patterns based on their anatomic sites of origin (*4*–*8*). Many of these differences were retained with ex vivo passage, which suggested a "transcriptional memory" of positional identities. Specifically, cells from each anatomic site had a unique combination of gene expression patterns that reflected the site's location along three developmental axes—anterior-posterior, proximal-distal, and cutaneous versus internal organs (*9*). Like a map based on longitudes and latitudes, cells reflect their anatomic origin by a digital combination of gene activities based on positional identity of cardinal axes from development. Indeed, position-specific gene expression programs are directly controlled by Hox proteins, which mediate anatomic-specific inductive effects (*10*).

Encoding Positional Identity: The Genome as a GPS Device

In a way, the genome functions as a global positioning system (GPS) device. Because ongoing *Hox* activity in adult cells is required for proper expression of genes denoting positional identity (*10*), a key challenge of cellular positional identity is keeping the right *Hox* genes turned on and off. In mammals, 39 *Hox* genes are clustered on four chromosomal loci, and their order along the chromosomes is roughly correlated with their spatial pattern of expression along the anterior-posterior and proximal-distal axes, a property termed colinearity (*11*). At the most fundamental level, positional identity is encoded in cells by the biochemical state of chromatin, the DNA-protein complex where *Hox* genes reside (*12*–*14*).

Eukaryotic DNA is wrapped around histone proteins, and this DNA-protein complex is subject to a large array of chemical modifications and protein interactions, which are believed to dictate whether the chromatin is accessible to gene activation or condensed for gene silencing (*15*). Comparison of chromatin states of cells from distinct anatomic sites has revealed salient features of how the chromatin states of *Hox* loci reflect the cells' positional identities (Fig. 1B) (*12*–*14*). Notably, in differentiated cells, the chromatin of the *Hox* loci can be programmed in either the ON or OFF state that extends across multiple *Hox* genes and intervening regulatory sequences, with the particular sequences in the ON or OFF state dependent on the positional identity of the cell along developmental axes. Extending our analogy of the map, each *Hox* locus can be considered a cardinal axis, like a longitude or latitude, and the chromatin state is the cellular mechanism for marking where on each axis the cell is located (Fig. 1B).

The ON state of the *Hox* loci is characterized by extensive occupancy of RNA polymerase II, intergenic transcription of long noncoding RNAs, and interaction with the trithorax group of proteins, which catalyze trimethylation of histone H3 on lysine 4. The OFF state is characterized by interaction with the Polycomb group of proteins, trimethylation of histone H3 on lysine 27 (H3K27me3), and also DNA methylation. Long noncoding RNAs, transcribed from within and outside of the *Hox* loci, also regulate chromatin states and gene expression by interfacing with trithorax and Polycomb complexes (*13*, *16*, *17*). Further, active *Hox* genes loop out from the remainder of the condensed chromosome territories in some tissues, which implies that the chromatin state of *Hox* genes can be reflected by—but does not strictly require—higher-order reorganization of the chromosomes (*18*). These chromatin states and corresponding *Hox* expression patterns of cells create a stable, cell-intrinsic memory of their positional identities. Chromatin patterns of tissues in vivo and of isolated cells ex vivo appear quite similar (*13*, *14*), even after extensive ex vivo passage (*10*), and *Hox* expression patterns are not altered by heterotypic cell-cell interactions (*10*). Nonetheless, trithorax, Polycomb, and long noncoding RNAs are continually required to reinforce positional identity (*13*, *19*). A recent finding that Polycomb proteins remain bound to DNA even during DNA replication provides another possible explanation for the persistence of the transcriptional memory of positional identity (*20*).

How did these chromatin states representing positional identities arise? It is known that enhancers, regulatory elements that stimulate gene expression, are located at both ends of the *Hox* loci. When pieces of the *Hox* loci were either deleted or inverted, the distance between the enhancers and each specific *Hox* gene was critical for the timing and spatial pattern of *Hox* expression (*21*, *22*). For instance, reducing the distance between a 3′ enhancer and a *HoxD* gene typically led to premature and more proximal expression of that *HoxD* gene. Surprisingly, the establishment of *Hox* chromatin states follows a more complex strategy. Inversion of a portion of the *HoxD* locus—moving several genes millions of bases away from their cognate enhancer—modestly

Howard Hughes Medical Institute and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. E-mail: howchang@stanford.edu

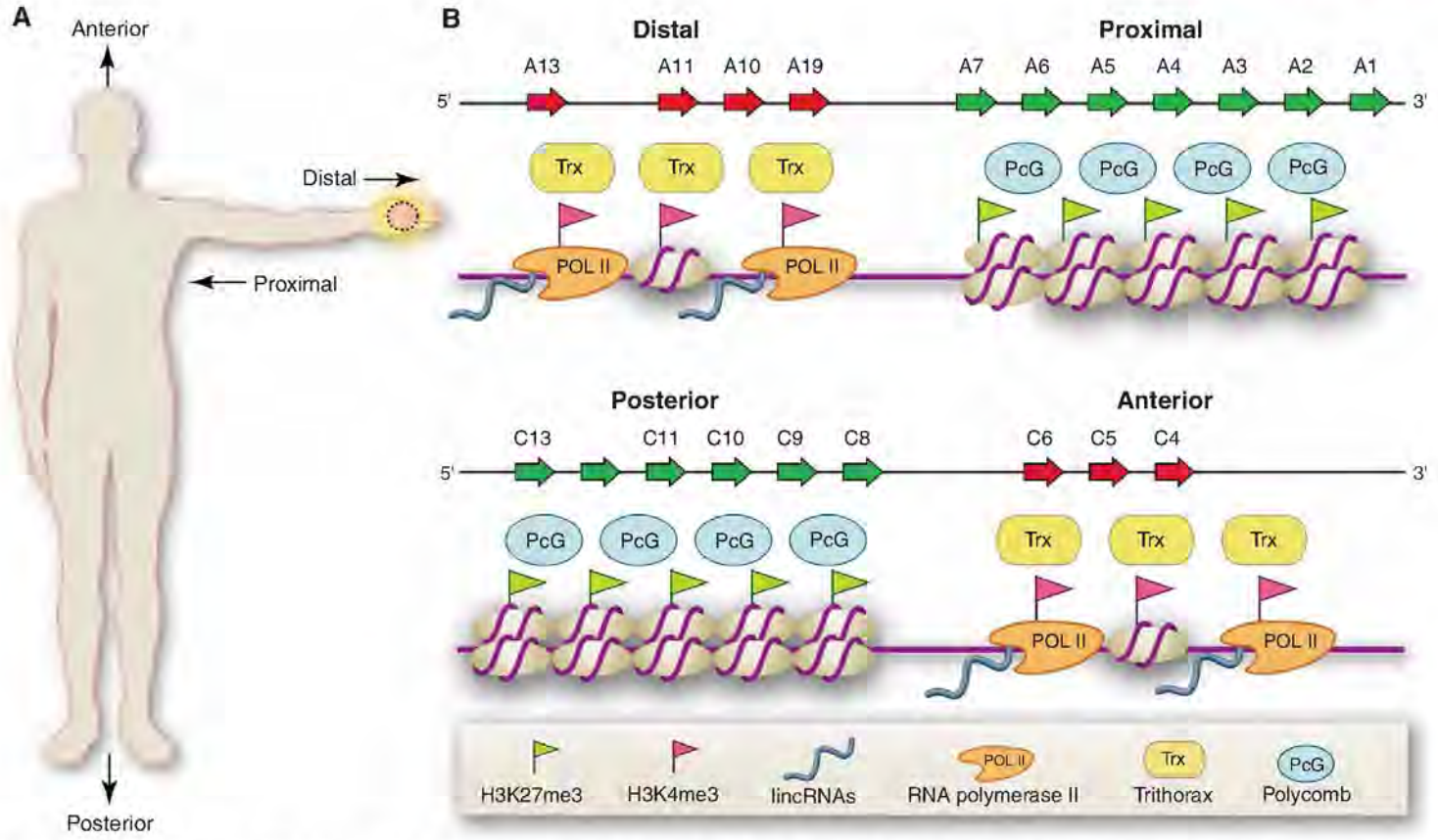


Fig. 1. The genome as a GPS device. (**A**) Developmental axes relative to human anatomy. The proximal-distal axis applies to all appendages such as upper limbs, lower limbs, and external genitalia. For instance, cells from the finger have anterior and distal positional identities. (**B**) Encoding positional identities in chromatin. Domains of open chromatin progress from 3′ to 5′ on all four of human *HOX* loci according to anterior-posterior position, but in different tissues. *HOXC* locus is strongly reflective of anterior-posterior identity in skin fibroblasts; whereas *HOXB* locus is active in cells from internal organs. Chromatin state on *HOXA* and *HOXD* loci demarcate proximal-distal identity. Thus, for cells in the finger, their positional identity is encoded by distal identity, which is reflected by active 5′ *HOXA* genes, and anterior identity is reflected by active 3′ *HOXC* genes. Active genes are shown in red; silent genes shown in green. Abbreviations: lincRNAs, long noncoding RNAs; H3K4me3, histone H3 lysine 4 trimethylation; H3K27me3, histone H3 lysine 27 trimethylation; Trx, trithorax group proteins; PcG, Polycomb group proteins.

affected chromatin marks associated with transcriptional activation in parallel with altered gene expression near the inversion breakpoint, but notably, had little effect on the ground state of the locus, characterized by broad occupancy of H3K27me3 (*14*). These and other findings imply the existence of parallel regulatory systems for encoding the transcriptional and chromatin states of *Hox* genes in relation to positional identity (*23*). The existence of multiple layers of regulatory sequences and extensive chromatin domains encompassing multiple *Hox* genes are likely to be driving forces for *Hox* genes to stay together and become more compact during vertebrate evolution (*24*).

Implications of Positional Memory

These insights into the cellular encoding of positional identity also point to gaps in our knowledge. How do cells of different positional identities sense each other and allocate cell fates to achieve proper proportion? This emergent property of large groups of cells certainly occurs in development (*25*) and implies a plasticity of positional identities and signaling mechanisms to link cell-cell interactions to changes in chromatin state. Plasticity of positional identity may be particularly important in wound repair when cells need to integrate their positional identity with the existing identities in the wound bed for proper regeneration (*26*). Cancer cells may also commandeer mechanisms involved in positional memory for their progression (*27*). Ultimately, these challenges need to be traced to understand how regulatory sequences specifically program chromatin states in the *Hox* loci and other master regulators of cell fate.

References and Notes

1. L. Wolpert, *Development* **107** (suppl.), 3 (1989).
2. L. Wolpert, *Dev. Genet.* **15**, 485 (1994).
3. N. Shubin, C. Tabin, S. Carroll, *Nature* **457**, 818 (2009).
4. H. Y. Chang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12877 (2002).
5. J.-T. Chi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10623 (2003).
6. J.-T. Chi *et al.*, *PLoS Genet.* **3**, 1770 (2007).
7. S. Gesta *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6676 (2006).
8. K. B. Ackema, J. Charité, *Stem Cells Dev.* **17**, 979 (2008).
9. J. L. Rinn, C. Bondre, H. B. Gladstone, P. O. Brown, H. Y. Chang, *PLoS Genet.* **2**, e119 (2006).
10. J. L. Rinn *et al.*, *Genes Dev.* **22**, 303 (2008).
11. M. Kmita, D. Duboule, *Science* **301**, 331 (2003).
12. B. E. Bernstein *et al.*, *Cell* **120**, 169 (2005).
13. J. L. Rinn *et al.*, *Cell* **129**, 1311 (2007).
14. N. Soshnikova, D. Duboule, *Science* **324**, 1320 (2009).
15. O. J. Rando, H. Y. Chang, *Annu. Rev. Biochem.* **78**, 245 (2009).
16. M. E. Dinger *et al.*, *Genome Res.* **18**, 1433 (2008).
17. A. M. Khalil *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11667 (2009).
18. E. Heard, W. Bickmore, *Curr. Opin. Cell Biol.* **19**, 311 (2007).
19. B. Schuettengruber, D. Chourrout, M. Vervoort, B. Leblanc, G. Cavalli, *Cell* **128**, 735 (2007).
20. N. J. Francis, N. E. Follmer, M. D. Simon, G. Aghia, J. D. Butler, *Cell* **137**, 110 (2009).
21. M. Kmita, N. Fraudeau, Y. Herault, D. Duboule, *Nature* **420**, 145 (2002).
22. B. Tarchini, D. Duboule, *Dev. Cell* **10**, 93 (2006).
23. C. Morey, N. R. Da Silva, M. Kmita, D. Duboule, W. A. Bickmore, *J. Cell Sci.* **121**, 571 (2008).
24. D. Duboule, *Development* **134**, 2549 (2007).
25. F. V. Mariani, G. R. Martin, *Nature* **423**, 319 (2003).
26. P. Leucht *et al.*, *Development* **135**, 2845 (2008).
27. A. Sparmann, M. van Lohuizen, *Nat. Rev. Cancer* **6**, 846 (2006).
28. I thank J. Helms and A. Oro for discussion and critiques. Supported by California Institute for Regenerative Medicine. H.Y.C. is an Early Career Scientist of the Howard Hughes Medical Institute.

10.1126/science.1175686

REVIEW

Actin, a Central Player in Cell Shape and Movement

Thomas D. Pollard[1]* and John A. Cooper[2]

The protein actin forms filaments that provide cells with mechanical support and driving forces for movement. Actin contributes to biological processes such as sensing environmental forces, internalizing membrane vesicles, moving over surfaces, and dividing the cell in two. These cellular activities are complex; they depend on interactions of actin monomers and filaments with numerous other proteins. Here, we present a summary of the key questions in the field and suggest how those questions might be answered. Understanding actin-based biological phenomena will depend on identifying the participating molecules and defining their molecular mechanisms. Comparisons of quantitative measurements of reactions in live cells with computer simulations of mathematical models will also help generate meaningful insights.

Life on Earth arose by divergent evolution from a common ancestor that lived ~3 billion years ago. Among its ~400 genes was the ancestral gene encoding actin (*1*). Actin and its bacterial counterparts polymerize into filaments (Fig. 1) that offer myriad advantages to cells. High cellular concentrations of actin make the protein one of the most abundant on earth. Actin is essential for the survival of most cells: Filaments provide internal mechanical support, tracks for movements of intracellular materials, and force to drive cell movements. Many modern species of prokaryotes use actin relatives to maintain asymmetrical shapes and to move DNA through the cytoplasm. Essentially all eukaryotes have genes for actin, and most have genes for myosin motor proteins that generate forces on actin filaments (*2*). In animals, actin filaments complement two other cytoskeletal polymers, microtubules and intermediate filaments (Fig. 1B).

Actin and myosin were discovered during the 1940s in muscle, where the two proteins comprise highly regular arrays of filaments that make up more than half of the total protein. Pioneering research on muscle established general principles that apply to actin assembly and function in all cells, including the mechanism used by myosin to produce force and movement from adenosine triphosphate (ATP) hydrolysis (*3*). Two decades later, actin and myosin were discovered in other cells (*4*, *5*), revealing that muscle filaments are a specialized example of a common cellular system. Subsequent research identified numerous proteins that regulate actin, analyzed their mechanisms of action, and linked the proteins to cellular processes.

Under physiological conditions, actin monomers (Fig. 2A) spontaneously polymerize into long, stable filaments (Fig. 2B) with a helical arrangement of subunits [for a review, see (*6*)]. Polymerization starts slowly, because small oligomers are very unstable, but once filaments have been created, actin polymerizes rapidly and almost completely. Actin filaments are polar, because the subunits in the filament all point in the same direction. One end of the filament grows much faster than the other end. Actin binds an adenine nucleotide [ATP or adenosine diphosphate (ADP)], and soon after assembly into filaments, actin hydrolyzes the terminal phosphate from the bound ATP and slowly dissociates the phosphate. Subtle changes in the structure of the actin subunits associated with these chemical reactions prepare ADP-actin filaments for disassembly by regulatory proteins.

Eukaryotic cells use >100 accessory proteins to maintain a pool of actin monomers, initiate polymerization, restrict the length of actin filaments, regulate the assembly and turnover of actin filaments, and cross-link filaments into networks or bundles (Fig. 2, C to F). The mechanisms forming new filaments include growing a branch on the side of an existing filament, severing a filament to create two new ends, or starting up a filament from monomers. Genes for most of these accessory proteins were in place about 1 billion years ago when the top branches formed on the phylogenetic tree, so amoebas, fungi, and animals share many molecular mechanisms that run their actin systems (*2*). Some species, such as the intestinal parasite *Giardia*, lack genes for myosin and many actin-binding proteins. These organisms may have diverged before these genes emerged (*7*), or they may have lost these genes, just as plants lost more than 200 genes required for the assembly of cilia and flagella (*8*).

Polymerization of actin filaments drives the crawling locomotion of eukaryotic cells, a characteristic feature of amoebas and animal cells (Fig. 3G). Actin polymerization also contributes to the internalization of membrane vesicles to control the composition of the cell membrane and the interface of the cell with the environment.

Interactions of myosin motor proteins with actin filaments (Fig. 2G) produce two types of movements. First, myosin generates force between actin filaments, producing contractions that pull up the rear of moving cells (Fig. 3G), pinch dividing cells in two (Fig. 3, D and E), and change cellular shapes to form tissues. A similar mechanism contracts muscle cells. Second, myosins associated with subcellular organelles and macromolecular complexes of proteins and RNA move these cargos along actin filaments over short distances (Fig. 3C). In budding yeast cells, which are small, actin filament tracks are responsible for distributing nearly all of the organelles and secretory vesicles to daughter cells before cell division. Movement of

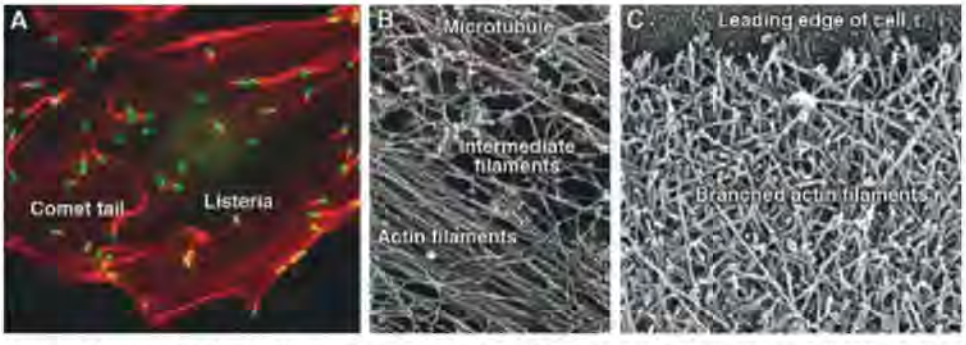


Fig. 1. Micrographs of actin filament structures in cells. **(A)** Fluorescence light micrograph of an animal epithelial cell grown in tissue culture and infected with a bacterium, *Listeria monocytogenes*. Actin filaments are red and bacteria are green. Actin bundles, called stress fibers, bridge sites of adhesion to the substrate. The bacteria use Arp2/3 complex to assemble comet tails for transport through the cytoplasm. **(B)** Electron micrograph of three types of cytoskeletal polymers in a cell permeabilized to release soluble components. After rapid freezing, the frozen water was sublimed away and cellular components were coated with platinum. Red colorization highlights a microtubule. A bundle of actin filaments and a network of intermediate filaments are labeled. **(C)** Electron micrograph of the network of branched actin filaments at the leading edge (top) of a motile keratocyte. The cell was grown in tissue culture, extracted to release soluble materials, dried, and coated with platinum. [All images are from (*38*). Sources are (A) Matthew Welch, University of California Berkeley, (B) John Heuser, Washington Univeristy St. Louis, and (C) Tatyana Svitkina, University of Pennsylvania, and Gary Borisy, Marine Biological Laboratory.]

[1]Department of Molecular, Cellular and Developmental Biology, Department of Molecular Biophysics and Biochemistry, and Department of Cell Biology, Yale University, Post Office Box 208103, New Haven, CT 06520–8103, USA. [2]Department of Cell Biology, Washington University, St. Louis, MO 63110, USA.

*To whom correspondence should be addressed: thomas.pollard@yale.edu

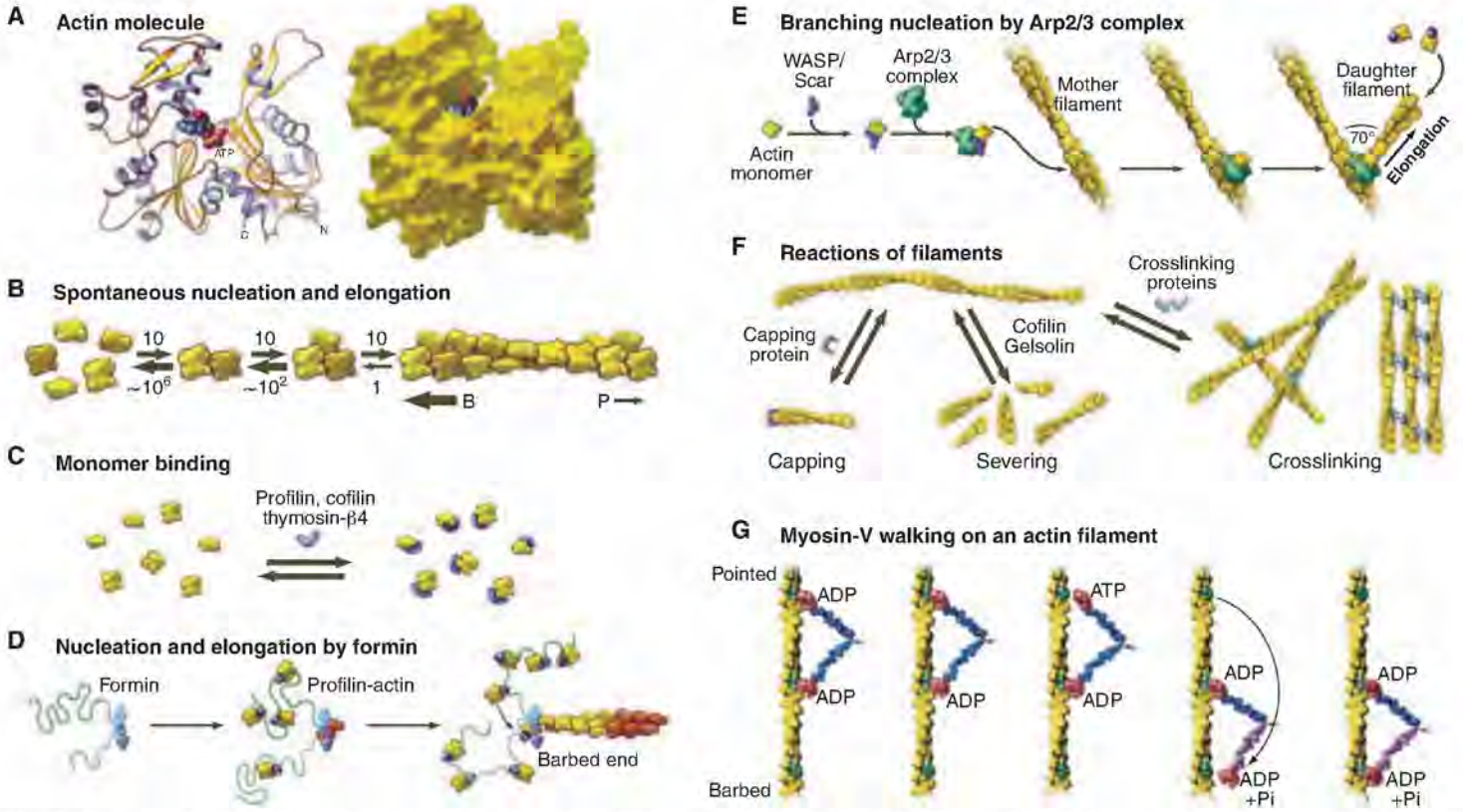


Fig. 2. Structures of actin and diagrams of fundamental reactions. (**A**) Ribbon and space-filling models of the actin molecule (pdb:1ATN). (**B**) Spontaneous nucleation and elongation. Dimers and trimers are unstable. Longer polymers grow rapidly at the barbed end (B) and slowly at the pointed end (P). (**C**) Actin monomer binding proteins. Thymosin-β4 blocks all assembly reactions; profilin promotes nucleotide exchange and inhibits pointed-end elongation and nucleation but not barbed-end elongation; cofilin inhibits nucleotide exchange and promotes nucleation. (**D**) Nucleation and elongation by formins. Formins initiate polymerization from free actin monomers and remain associated with the growing barbed end. Profilin-actin binds to formin and transfers actin onto the barbed end of the filament. (**E**) Nucleation by Arp2/3 complex. Nucleation-promoting factors such as WASp bind an actin monomer and Arp2/3 complex. Binding to the side of a filament completes activation, and the barbed end of the daughter filament grows from Arp2/3 complex. (**F**) Reactions of actin filaments. Capping proteins bind to and block barbed ends; cofilin and gelsolin sever filaments; cross-linking proteins assemble networks and bundles of actin filaments. (**G**) Myosin motors, such as myosin V, use cycles of ATP hydrolysis to walk along actin filaments, generally toward the barbed end. Redrawn from images in (*38*).

cargo over longer distances in many other cell types involves microtubules and their motors.

Questions to Ask

First, researchers need a complete list of parts to understand how the actin system or any biological system works. Do we know all the parts of the actin system? We know many, but the list is far from complete. Second, which parts interact with other parts? Third, how does this system of connected parts work as a whole? Most biological systems are so complicated that their operations are not intuitively obvious, so mathematical models and simulations are needed to connect hypotheses with experimental observations (*9*).

Examples of Biological Processes That Depend on Actin

Making connections between molecules and biology can be challenging. On one hand, any cellular process depends on many different proteins. On the other hand, a given molecule will contribute many processes, as shown by doing a PubMed search on any protein mentioned here. For example, recent publications provide evidence that cofilin participates in cancer, embryonic development, HIV infection, pathfinding by nerve cell axons, learning and memory, programmed cell death, Alzheimer's disease, traffic of intracellular membranes, mitosis, cytokinesis, tight junctions, and immune reactions of T-lymphocytes. In each of the following examples, the inventories of participating molecules are more advanced than our understanding of the molecular mechanisms or the operations of these processes at the system level.

Actin filaments as part of the cytoskeleton. The protein polymers forming the cytoskeleton are responsible for the mechanical properties and shapes of cells, which are often critical to their functions. If the membranes of a human cell are dissolved away to release soluble components, a ghostlike meshwork of cytoskeletal polymers remains (Fig. 1B) (*10*). The polymers include actin filaments, microtubules, and intermediate filaments in various proportions and geometries. Actin filaments provide mechanical structure and motility for amoeboid and animal cells. Microtubules are responsible for separating chromosomes and long-range transport of large particles in all eukaryotes. Intermediate filaments in vertebrates function as intracellular ligaments and tendons to resist mechanical forces.

Interactions among the three cytoskeletal polymers reinforce the cytoskeleton, although some cross-linking proteins exchange rapidly and the polymers themselves turn over on time scales of seconds to minutes. These features give the cytoplasm useful properties, such as being stiff when deformed rapidly and malleable when deformed slowly. Even the cells of plants and fungi, despite being encased in a cell wall, use cytoskeletal polymers to direct the shape of their compartments (*11*). In addition, the cytoskeleton is part of a system that senses both external forces applied to the cell and the mechanical properties of the cell's environment. This system can influence diverse aspects of cell function, including gene expression and differentiation (*12*).

Actin patches and endocytosis. Actin filaments assemble at sites of plasma membrane internalization in budding and fission yeast (*13*, *14*). In these "actin patches," filaments assemble de novo, provide force to form and internalize an endocytic vesicle from the plasma membrane, and then disassemble in a process that is self-limited

in space and time (Fig. 3A). Although powerful molecular and genetic tools have identified ~30 to 50 participating proteins, the parts list still appears to be incomplete. Actin is associated with endocytosis in many cells besides yeast, with an apparently similar set of molecular players (*15*).

The process of endocytosis begins at multiple independent sites, with the spontaneous assembly of membrane proteins along with clathrin and adaptor proteins. Next to be recruited are proteins including WASp family proteins and certain class I myosins that bind to and/or activate Arp2/3 complex, which creates new filaments as branches on older filaments (Fig. 2E). The source of the very first filaments remains unclear. Capping protein limits the growth of actin filaments, and filaments are linked together along their sides by fimbrin, among other proteins.

Although assembly of such a network of filaments alone can create force sufficient to deform a membrane, specialized proteins associate with membrane to induce curvature. The density of actin filaments decreases rapidly as the endocytic vesicle moves into the cytoplasm, a process that depends on the filament-severing protein cofilin, perhaps aided by proteins Aip1 and coronin (*16*). Although actin patches are one of the best characterized actin systems, our understanding of these reactions is limited, and some apparently contradictory observations exist, illustrating just how little we know about the process.

Bacterial comet tails. After invading a eukaryotic cell, some bacteria usurp cellular proteins to assemble a comet tail of actin filaments for propulsion through the cytoplasm (Fig. 1A and Fig. 3, B and F). Nucleation-promoting proteins on the surface of the bacterium recruit Arp2/3 complex to polymerize actin filaments. Growth of those filaments pushes the bacterium forward. The whole process can be reconstituted with the bacterial nucleation-promoting protein on the surface of a bead or lipid vesicle in a solution with purified actin, profilin, Arp2/3 complex, a capping protein, and the severing protein cofilin (*17*) and then simulated with a computer model (*18*).

Cytokinesis. The physical separation of two daughter cells is the last step in the cell cycle (Fig. 3, D and E). Amoebas, fungi, and animals use a contractile ring of actin filaments and myosin II to pinch themselves in two. Myosin II polymerizes into bipolar filaments, which can produce a contraction by pulling actin filaments together. Multicellular animals adapted the contractile ring machinery in specialized cells that evolved into muscle. Organisms on the other branch of the tree (including algae, plants, and ciliates) lack myosin II (*2*), so cytokinesis depends primarily on membrane fusion (plants) or on mechanisms that are not understood. Remarkably, prokaryotes use a protein related to the microtubule subunit tubulin to assemble a ring of filaments that pinches these cells in two, much like a contractile ring but without the obvious participation of a motor protein (*19*).

Having started with a common genetic toolbox for cytokinesis, one expects amoebas, fungi, and animals to use similar mechanisms for cytokinesis (*20*). Nevertheless, the general principles are still cloudy for two reasons. First, the process is very complicated, involving the products of >130 genes in fission yeast, a molecular inventory shared at least in part by other less well-characterized organisms. Second, some cells emphasize certain features more than other cells, because certain cytokinesis genes were added or lost from various lineages over the past billion years.

Successful cytokinesis depends on (i) placing the cleavage furrow in the right place between the separated chromosomes, (ii) assembling, constricting, and disassembling the contractile ring, and (iii) fusing the plasma membrane between the daughter cells. In animal cells, the information to place the cleavage furrow comes from the mitotic spindle, part from the poles of the spindle and part from the center of the spindle, where the chromosomes initially congregate. The cleavage site around the equator is marked with active signaling proteins called Rho GTPases (guanosine triphosphatase) (*21*). The mechanisms linking these GTPases to the assembly of the contractile ring are still being investigated. In fission yeast, negative signals from the ends of the cell and positive signals from the centrally placed nucleus localize contractile ring precursors called nodes to the middle of the cell (*20*). Nodes accumulate myosin II along with a formin protein that nucleates the growth of actin filaments. Computer simulations established the feasibility of one hypothesis for the assembly of the ring: Myosin molecules capture actin filaments growing randomly from adjacent nodes and pull the nodes together into a ring over 10 min (*22*). Contractile ring assembly is less well understood in other cells (*23*).

After the mitotic apparatus separates the two daughter nuclei, the contractile ring constricts, pulling the cell membrane into a cleavage furrow. Remarkably, the proteins of the contractile ring disperse as it constricts. Membrane fusion resolves the membranes of the two daughter cells (*24*).

Actin cables and organelle transport. Many, perhaps all, eukaryotic cells use myosin motors to transport organelles along actin filaments (Fig. 3C). Budding yeast replicate by directing secretion of cell-wall materials to grow a bud from a particular location on the plasma membrane of a mother cell. At a predetermined bud site, molecular polarity cues activate formins to nucleate actin filaments. A formin remains associated with each fast-growing barbed end to promote elongation (at 200 subunits per second) and prevent capping. The uniformly polarized filaments form bundles that serve as tracks for the movement of organelles (*25*). Class V myosin "walks" toward the barbed ends of these filaments (Fig. 2G) and moves secretory vesicles and intracellular organelles to the bud (*26*). Tropomyosin lies along the actin filaments to stabilize the bundles; it may also influence the action of the myosin motor. Myosin V also moves certain mRNAs on cables into the daughter, to influence cell fate and fitness.

Fission yeast (*27*) and plant cells (*28*) also depend on formins to assemble uniformly polarized actin filament cables as tracks for transport of materials for polarized growth. In animal cells and in elongated fungal hyphae, long-range movements depend largely on microtubules, and actin filaments do not appear to be organized into cables of uniform polarity. However, myosins coordinate with microtubule motors to move organelles over short distances along the actin filaments (*29*).

Cellular motility. Actin filaments are essential for cell locomotion, a defining feature of animal cells (*30*). For example, cells of the immune system migrate to search and destroy pathogens or cancer cells. During development of animal embryos, some cells move from one location in the body to another, by crawling between neighboring cells and through the extracellular matrix. Cancer cells use similar mechanisms to spread through the body. Nerve cells provide spectacular examples of both cell migration and cell process extension. Neurons destined to control the intestine migrate large distances as neural crest cells during development (*31*), nerve cells grow processes up to 1 m long to find their targets, and the human brain has 1.5 million km of such cellular processes.

Assembly of actin filaments from their monomeric subunits can suffice to change the shape of the cell and produce a protrusion, which is often the first step in cell locomotion. Arp2/3 complex assembles a dense network of short, branched actin filaments (Figs. 1C and 3B) that grow in successive generations like the twigs of a bush (*32*). Each filament can produce piconewton forces (*33*), allowing the front end of cells to move at rates up to about 1 μm per second (*34*). Most filaments are capped before growing longer than 0.5 μm; longer filaments would presumably buckle under force. Computer simulations reveal that the components have the capacity for self-organization into networks (*35*).

A short distance behind the leading edge, the network of branched filaments turns over within a few seconds, replaced by one composed of longer unbranched filaments (*34*). Formins are logical candidates to help remodel the network in this manner, because formins are known to create long unbranched actin filaments in thin protrusions called filopodia (*36*). The relative contributions of formins and Arp2/3 complex to motility have been difficult to sort out, in part because Arp2/3 complex is essential for viability and active at very low concentrations (making depletion experiments difficult). Specific pharmacologic inhibitors for Arp2/3 complex only recently became available (*37*). During cell locomotion, myosin interacts with actin filaments to pull up the rear of the cell, working as structural elements with myosin motor proteins.

Conclusions

The big questions for the actin system are not so much about "what happens" but, rather, "how

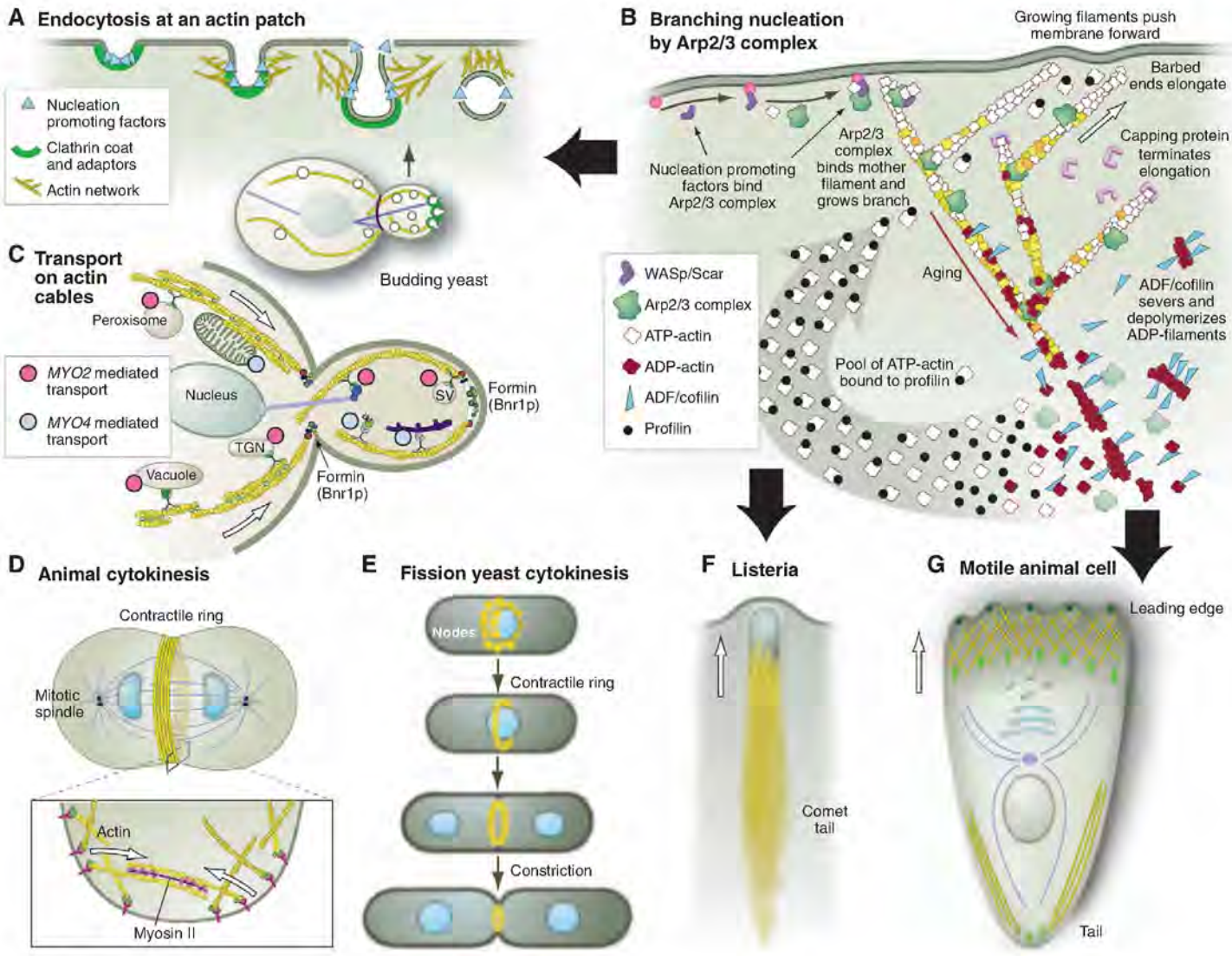


Fig. 3. Cartoons showing actin-based movements. (**A**) Clathrin-mediated endocytosis at fungal actin patches. (**B**) Formation of branched filament networks nucleated by Arp2/3 complex, which is used for three of these types of movement, those depicted in (A), (F), and (G). (**C**) Transport of membrane-bound vesicles, organelles, and RNAs from the mother cell to the daughter cell by class V myosins in budding yeast. (**D**) Cytokinesis in animal cells by constriction of a contractile ring of actin filaments and myosin II. (**E**) Cytokinesis in fission yeast. The contractile ring of actin filaments and myosin II forms by condensation of nodes. (**F**) The *Listeria* bacterium stimulates the assembly of an actin filament comet tail to push it through the cytoplasm of a host animal cell. (**G**) Locomotion of an animal cell by assembly of actin filaments at the leading edge and retraction of the tail. Redrawn from (*13*)(A), (*32*)(B), (*25*)(C), (*38*)(D and E), (*18*)(F), and (*30*)(G).

these systems work" at the molecular level. Given the common origin of the genes for the actin system, evolution should be able to help sort out the complicated mechanisms. Continuing focus on tractable model systems should help to establish the molecular basis for each actin-based function and general principles that apply more broadly. Research with systems-level genomics approaches based on genetic and physical interactions in model organisms continues to increase the parts list and to reveal new interactions. The missing rate and equilibrium constants required to trace pathways and explain how molecules work together still need to be defined. The regulation of actin by signaling mechanisms and the interaction of actin with other cellular systems, such as membranes, requires further attention.

Technical advances should prove critical. For example, we can now image the behavior of single molecules over time, in vitro and in cells, and we can reconstitute complex processes using mixtures of purified components. Furthermore, advances in light and electron microscopy allow for nanometer-level localization of protein components and for measurement of global and local concentrations of molecules inside living cells. These advances should prove critical to advance our understanding of these actin systems, especially how and where filaments are created and assembled into networks of varying geometry.

The field has only recently started to create mathematical models at the microscopic, mesoscopic, and macroscopic scales. Still, great progress has been made in several areas, such as understanding how the actin filaments in a protrusion assemble and create force on the plasma membrane (*9*, *18*, *35*). Quantitative measurements in live cells aided by genetics, specific drugs, and depletion strategies should provide the data required to test hypotheses embodied in mathematical models.

References and Notes

1. H. P. Erickson, *Bioessays* **29**, 668 (2007).
2. T. A. Richards, T. Cavalier-Smith, *Nature* **436**, 1113 (2005).

3. M. A. Geeves, K. C. Holmes, *Adv. Protein Chem.* **71**, 161 (2005).
4. S. Hatano, F. Oosawa, *Biochim. Biophys. Acta* **127**, 488 (1966).
5. M. R. Adelman, E. W. Taylor, *Biochemistry* **8**, 4964 (1969).
6. T. D. Pollard, *Annu. Rev. Biophys. Biomol. Struct.* **36**, 451 (2007).
7. H. G. Morrison *et al.*, *Science* **317**, 1921 (2007).
8. D. Lang, A. D. Zimmer, S. A. Rensing, R. Reski, *Trends Plant Sci.* **10**, 542 (2008).
9. A. Mogilner, R. Wollman, W. F. Marshall, *Dev. Cell* **11**, 279 (2006).
10. J. E. Heuser, M. W. Kirschner, *J. Cell Biol.* **86**, 212 (1980).
11. P. J. Hussey, T. Ketelaar, M. J. Deeks, *Annu. Rev. Plant Biol.* **57**, 109 (2006).
12. D. E. Discher, D. J. Mooney, P. W. Zandstra, *Science* **324**, 1673 (2009).
13. B. J. Galletta, J. A. Cooper, *Curr. Opin. Cell Biol.* **21**, 20 (2009).
14. M. Kaksonen, C. P. Toret, D. G. Drubin, *Nat. Rev. Mol. Cell Biol.* **7**, 404 (2006).
15. G. J. Doherty, H. T. McMahon, *Annu. Rev. Biochem.* **78**, 857 (2009).
16. W. M. Brieher, H. Y. Kueh, B. A. Ballif, T. J. Mitchison, *J. Cell Biol.* **175**, 315 (2006).
17. T. P. Loisel, R. Boujemaa, D. Pantaloni, M. F. Carlier, *Nature* **401**, 613 (1999).
18. J. B. Alberts, G. M. Odell, *PLoS Biol.* **2**, e412 (2004).
19. M. Osawa, D. E. Anderson, H. P. Erickson, *Science* **320**, 792 (2008).
20. S. Oliferenko, T. G. Chew, M. K. Balasubramanian, *Genes Dev.* **23**, 660 (2009).
21. A. L. Miller, W. M. Bement, *Nat. Cell Biol.* **11**, 71 (2009).
22. D. Vavylonis, J. Q. Wu, S. Hao, B. O'Shaughnessy, T. D. Pollard, *Science* **319**, 97 (2008).
23. M. Zhou, Y. L. Wang, *Mol. Biol. Cell* **19**, 318 (2008).
24. G. Montagnac, A. Echard, P. Chavrier, *Curr. Opin. Cell Biol.* **20**, 454 (2008).
25. D. Pruyne, A. Legesse-Miller, L. Gao, Y. Dong, A. Bretscher, *Annu. Rev. Cell Dev. Biol.* **20**, 559 (2004).
26. A. Fagarasanu, R. A. Rachubinski, *Curr. Opin. Microbiol.* **10**, 528 (2007).
27. S. G. Martin, F. Chang, *Curr. Biol.* **16**, 1161 (2006).
28. L. Vidali *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13341 (2009).
29. V. Levi, A. S. Serpinskaya, E. Gratton, V. Gelfand, *Biophys. J.* **90**, 318 (2006).
30. A. J. Ridley *et al.*, *Science* **302**, 1704 (2003).
31. R. B. Anderson, D. F. Newgreen, H. M. Young, *Adv. Exp. Med. Biol.* **589**, 181 (2006).
32. T. D. Pollard, G. G. Borisy, *Cell* **112**, 453 (2003).
33. D. R. Kovar, T. D. Pollard, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14725 (2004).
34. T. M. Svitkina, A. B. Verkhovsky, K. M. McQuade, G. G. Borisy, *J. Cell Biol.* **139**, 397 (1997).
35. T. E. Schaus, E. W. Taylor, G. G. Borisy, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7086 (2007).
36. C. Yang *et al.*, *PLoS Biol.* **5**, e317 (2007).
37. B. J. Nolen *et al.*, *Nature* **460**, 1031 (2009).
38. T. D. Pollard, W. C. Earnshaw, *Cell Biology* (Saunders, New York, 2007).

39. This work was supported by NIH research grants GM026132, GM026338, GM066311, GM38542, and GM47337.

10.1126/science.1175862

REVIEW

Subcellular mRNA Localization in Animal Cells and Why It Matters

Christine E. Holt[1] and Simon L. Bullock[2]*

Subcellular localization of messenger RNAs (mRNAs) can give precise control over where protein products are synthesized and operate. However, just 10 years ago many in the broader cell biology community would have considered this a specialized mechanism restricted to a very small fraction of transcripts. Since then, it has become clear that subcellular targeting of mRNAs is prevalent, and there is mounting evidence for central roles for this process in many cellular events. Here, we review current knowledge of the mechanisms and functions of mRNA localization in animal cells.

The asymmetric distribution of specific mRNAs in the cytoplasm was first visualized in the early 1980s, when in situ hybridization techniques were used to detect β-*actin* mRNA in ascidian embryos (*1*). The discovery of differential localization of transcripts encoding cytoskeletal proteins in cultured chicken cells soon gave further prominence to this phenomenon (*2*). Subsequent studies demonstrated that asymmetric mRNA localization contributes to the targeting of diverse types of protein products.

In recent years, the advent of high-throughput approaches has revealed that mRNA localization is much more common than previously assumed. Of expressed mRNA species, 70% were classified as asymmetrically distributed in a large-scale fluorescent in situ hybridization screen in early *Drosophila* embryos (*3*). In addition, large numbers of vertebrate mRNAs are specifically enriched in protrusions of migrating fibroblasts, in neuronal processes, or on spindles (table S1). Thus, mRNA localization has a prominent role in the spatial regulation of gene activity. Here, we provide an overview of the mechanisms and functions of mRNA localization in animal cells. Readers are referred elsewhere for entry points into the seminal work on mRNA localization in fungi and plants (*4*, *5*).

[1]Department of Physiology, Development, and Neuroscience, University of Cambridge, Cambridge CB2 3DY, UK. [2]Cell Biology Division, Medical Research Council (MRC) Laboratory of Molecular Biology, Cambridge CB2 0QH, UK.

*To whom correspondence should be addressed. E-mail: sbullock@mrc-lmb.cam.ac.uk

Mechanisms of mRNA Localization: Illuminating a Multi-Step Process

Four mechanisms are thought to contribute to subcellular localization of specific mRNAs after their transcription: (i) vectorial export from nuclei, (ii) localized protection from degradation, (iii) polarized active transport on the cytoskeleton by using molecular motors, and (iv) localized anchorage. With the exception of vectorial nuclear export, all of these mechanisms are known to contribute to mRNA sorting in animal cells. Combinations of these mechanisms can also be used to localize a single mRNA species.

Protection of mRNAs from degradation (Fig. 1A) plays a crucial role in restricting mRNAs to the germ plasm in *Drosophila* and zebrafish embryos, often in conjunction with local entrapment of transcripts (*6*–*8*). There is also evidence, from the sea slug *Aplysia*, that mRNAs in neuronal processes can be selectively stabilized by interaction with their targets (*9*). However, the molecular mechanisms that locally protect specific messages remain unknown.

Motor-based transport (Fig. 1B) appears to be the predominant mechanism for the localization of mRNAs in animal cells probably because it provides the most rapid method for long-distance translocation of large ribonucleoprotein (RNP) particles through the crowded cellular environment. Live cell-imaging studies in recent years—involving the injection of in vitro synthesized fluorescent mRNAs or labeling transcripts by means of tethering multiple fluorescent proteins—have provided compelling evidence that mRNAs can control their own sorting by recruiting more than one kind of motor and even modulating motor properties.

For instance, in mammalian oligodendrocytes and hippocampal neurons, as well as in *Drosophila* embryos, mRNAs are bound to microtubule-based motor complexes that rapidly switch between bouts of motion in the minus- and plus-end directions (*10*–*12*). Specific mRNAs appear to control net sorting by increasing the relative frequency of movement in one direction through the recruitment of factors that modulate the activities of simultaneously bound opposite polarity motors (*11*).

In the case of delivery of *oskar* mRNA from the nurse cells to the posterior pole of the *Drosophila* oocyte, the frequency of microtubule-based movement in the minus-end and plus-end directions is also altered by specific components of messenger RNPs (mRNPs) (*13*). However, it appears that this comprises sequential, rather than rapidly switching, actions of motors. Localization of *oskar* culminates in a biased walk along a weakly polarized cytoskeleton—driven by the plus end–directed motor kinesin-1—to anchorage sites at the posterior pole (*13*). Vegetal localization of mRNAs in *Xenopus* oocytes may also be based on similar principles, although in this case the concerted action of kinesin-1 and kinesin-2 is crucial (*14*).

Some mRNAs, as is the case for other cellular cargoes, may simultaneously associate with actin-

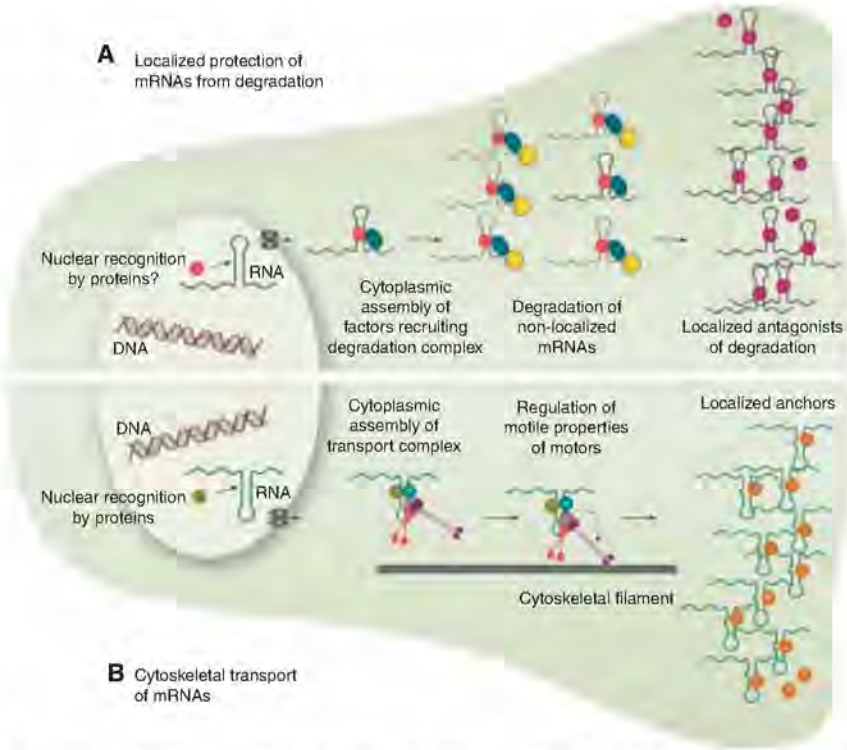


Fig. 1. mRNA localization is a multi-step process. Shown is an illustration of two stylized cells, depicting mechanisms that can contribute to mRNA localization. (**A**) Protection of mRNAs from degradation. Red, nuclear RNA recognition factor; dark blue, cytoplasmic RNA recognition factor; yellow, ribonuclease; purple, agonist of degradation. (**B**) Motor-based transport. Green, nuclear RNA recognition factor; light blue and light gray, cytoplasmic RNA recognition factor; red and purple, molecular motors; orange, anchorage factor. In reality, different combinations of these mechanisms may be used to localize a single mRNA species in the same cell.

and microtubule-based motors, allowing transport to be fine-tuned by switching between different types of cytoskeletal tracks (*15*). Transcripts may also influence the choice of subsets of microtubules by motors. This mechanism has been proposed to contribute to the delivery of *gurken* and *bicoid* mRNAs to the dorso-anterior and anterior regions of the *Drosophila* oocyte, respectively, by the minus end–directed motor dynein and could conceivably be based on differential posttranslational modification of microtubules (*16*, *17*).

Although our understanding of transport mechanisms is increasing, relatively little is known about the processes that contribute to mRNA anchorage. Long-distance transport of mRNPs on microtubules can be followed by transfer to the actin cytoskeleton at the cortex, with entrapment facilitated by the dense network of filaments or associated proteins (*18*, *19*). In other cases, microtubule-based motors may act directly as anchors (*20*) or lead to steady-state mRNA localization through continual active transport (*21*).

Thus, it appears that multiple binding sites within mRNAs recruit combinations of trans-acting factors that regulate the association and activities of different molecular motors as well as mediating interplay with anchorage complexes and translational regulators (see below). Even uniformly distributed mRNAs can be transported to some extent by motors, presumably to facilitate their exploration of space (*11*, *22*, *23*). A key challenge for the future is to understand how the information within asymmetrically localized transcripts is decoded.

Recognition of Localizing mRNAs

Cis-acting RNA localization signals and interplay with translation. The cis-acting elements that mediate asymmetric localization of specific transcripts are referred to as RNA localization signals or zip codes. Depending on the nature of the trans-acting factor they bind, these elements can consist of single-stranded stretches or double-stranded stem loops (*24*). Characterizing these latter types of elements is taxing because recognition may be based on a three-dimensional structure. This is probably the case during transport of several mRNAs toward the minus ends of microtubules in *Drosophila*, where stem loops with relatively little in common at the primary sequence level are recognized by the same RNA-binding protein, Egalitarian (*25*).

Localization signals are typically found within untranslated regions of messages, where they can evolve without the constraints of retaining protein-coding sequences. In cases in which signals are found within coding sequences, their secondary structure may play a role in antagonizing the translational machinery during the mRNA localization process (*26*). Protein production is more commonly silenced during translocation by the recruitment of translational repressors (*27*). In some instances, initiation of protein synthesis at the target site is mediated by the interplay between localized translational derepressors and proteins that bind localization signals. An elegant example of this involves the phosphorylation of the β-*actin* zip code–binding protein ZBP1 by the localized activity of the kinase Src at the cell periphery (*28*). This leads to dissociation of ZBP1 from the transcript at the leading edge of migrating cells, allowing access to the translation initiation machinery.

Trans-acting factors and the assembly of mRNPs. A large number of proteins have been identified with direct roles in mRNA localization complexes. To what extent this reflects discrete pathways at work or functionally related mRNPs containing multiple proteins remains unclear. This latter scenario will be at least part of the story because there are several reports of combinations of well-characterized RNA-binding proteins, such as ZBP1, Staufen, and fragile X mental retardation protein (FMRP), being found in the same complexes. Interactions of the same RNA with multiple trans-acting factors gives scope for redundancy, which may partly explain the difficulties in identifying the molecular links between localizing transcripts and motors in animal cells. However, a complete link has recently been uncovered between mRNA localization signals and dynein in *Drosophila* (*25*), providing an opportunity to probe the molecular details of the assembly and operation of a model RNA: motor complex.

Where in the cell are mRNAs earmarked for transport? In many instances, localizing transcripts are first recognized in the nucleus. This is the case for β-*actin* transcripts in chicken fibroblasts, in which the cotranscriptional recruitment of the ZBP2 protein facilitates binding of ZBP1 to the mRNA and its subsequent targeting behind the leading edge (*29*). It has also been revealed, from elegant studies of *Vg1* localization in *Xenopus* oocytes, that important RNA:protein interactions formed in the nucleus can be remodeled in the cytoplasm (*30*), and such events may regulate transitions between critical steps in localization processes. Nuclear history also plays an essential role in cytoplasmic localization of *oskar* mRNA. Deposition of the multicomponent exon junction complex (EJC) during splicing is essential for the translocation of this transcript to the posterior of the *Drosophila* oocyte (*31*), possibly by facilitating switching of the predominant motor activity from dynein to kinesin-1 (*13*). It will be fascinating to discover how the EJC regulates these motors at the molecular level, especially because components of this complex

have been implicated in the localization of functionally important mRNAs within mammalian neurons (*32*).

Functions of mRNA Localization: Cell Polarity and Local Response to Extrinsic Cues

There are several a priori reasons why localizing an mRNA could be advantageous over targeting the protein product directly: (i) increased cost effectiveness because of the production of multiple protein copies from single localized mRNA molecules, (ii) preventing proteins from acting ectopically during translocation, (iii) facilitating the assembly of macromolecular protein complexes by producing a high local concentration of mRNA molecules in microdomains, (iv) distinct properties of newly synthesized proteins, and (v) decentralizing the control of gene expression by permitting local control of translation in response to extrinsic cues. Below, we introduce specific examples that illustrate the importance of asymmetric mRNA localization in key biological processes (see also Fig. 2).

Establishing Embryonic Organization

In *Drosophila*, the differential localization of maternal mRNAs plays a major role in establishing and patterning the dorsal-ventral and anterior-posterior body axes as well as in germ cell specification (table S2). During *Xenopus* development, localization pathways exist in early and late oogenesis that culminate in the vegetal accumulation of transcripts that are important for germline development and patterning of the mesoderm and endoderm (*33*). Differentially localized maternal mRNAs have also been found in ascidians and cnidarians, and many of these transcripts encode proteins with known roles in embryonic patterning (*34*, *35*). Thus, the localization of maternal mRNAs appears to be widely used to establish embryonic organization.

In mammals, an obligatory requirement for localized mRNA determinants in the egg appears to be ruled out by the developmental lability of the early embryo. However, the recent report of apical localization of the message encoding the Cdx2 transcription factor in 8- to 16-cell embryos raises the possibility that mRNA sorting facilitates asymmetric cell fate decisions at later stages (*36*). A function for mRNA localization in influencing embryonic cell lineage choices is also supported by the differential inheritance of messages encoding developmental regulators in snail blastomeres, which is driven by a remarkable process of transcript enrichment at one of the two centrosomes (*37*).

Neurons: mRNA Localization on Demand

The critical importance of posttranscriptional regulation in neurons is illustrated by the high degree of autonomy exhibited by neuronal processes, which often extend great distances from the cell body. This autonomy permits rapid local responses to extrinsic cues and is manifest in the ability of axons and dendrites, respectively, to navigate to guidance cues and undergo certain forms of synaptic plasticity after removal of the soma. It has become increasingly evident that this "decentralization" involves the selective localization and translation of subsets of mRNAs in neuronal processes in response to external stimuli (Fig. 3).

Synapse formation and plasticity. In mammalian hippocampal neurons, strong synaptic activation is accompanied by transcription of the *Arc* gene and rapid trafficking of its mRNA to dendrites, where it localizes selectively to active synaptic sites (*38*). Arc is required for the consolidation of long-term potentiation (LTP), a form of persistent synaptic change, most likely through its ability to regulate actin dynamics and to modulate dendritic spine morphogenesis (*39*). Direct evidence for a requirement for mRNA localization in synaptic plasticity comes from studies of *CamKIIα*. Disruption of dendritic targeting of this mRNA in mice, by replacing the 3′ untranslated region of the endogenous gene with one from a nonlocalizing transcript, impairs LTP and long-term memory (*40*).

mRNA localization is also important for the establishment of synapses. In *Aplysia* sensory neurons, contact with a target motor neuron triggers rapid local concentration of the neuropeptide-encoding *sensorin* mRNA into synapses (*41*). Synaptogenesis is disrupted when *sensorin* mRNA levels are acutely reduced, even before the total concentration of the protein is altered. This indicates not only that mRNA localization is important but that newly synthesized Sensorin could have properties distinct from those of older protein copies. Consistent with an important role for nascent Sensorin, its translation is spatially restricted to active synapses in a stimulus-specific manner (*42*).

Cue-induced mRNA localization in axons. Growing axons navigate in the developing brain using attractive and repulsive cues that stimulate changes in growth and directional steering. β-*actin* mRNA is abundant in *Xenopus* growth cones and is rapidly recruited to the near-stimulus side in response to an attractive gradient (*43*, *44*). Attractive turning is abolished through the specific inhibition of local β-*actin* mRNA translation or disruption of the interaction of Vg1RBP (the *Xenopus* ZBP1 ortholog) with the zip code (*43*, *44*). The picture that emerges of localized translation of mRNAs underlying directionally responsive cell protrusions is strikingly similar to the situation in chicken fibroblasts (see below) and suggests that common mechanisms span the two systems (Fig. 3A).

"On site" versus "distant site" for action. Proteins synthesized from spatially localized mRNAs commonly act "on site." But this is not always the case. The mRNA encoding the transcription factor cyclic adenosine monophosphate (cAMP) response element–binding protein (CREB), which promotes cell survival in dorsal root ganglia neurons, can be translated locally in axons in response to nerve growth factor (*45*). The nascent CREB protein undergoes retrograde transport to the nucleus, where it activates the transcription of target genes. There is evidence that the phosphorylation status of CREB differs depending on its site of translation (*45*), which raises the intriguing possibility that local translation of process-targeted mRNAs controls gene expression in response to distal experience.

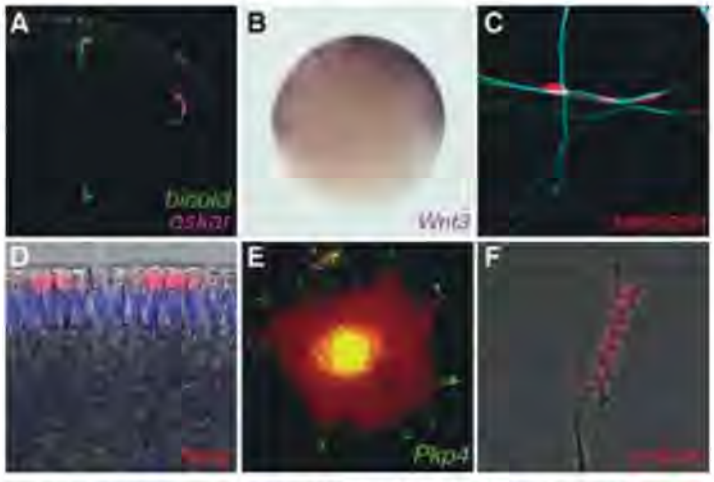


Fig. 2. Examples of asymmetrically localized mRNAs. (**A**) Differential localization of mRNA determinants within the *Drosophila* oocyte. (**B**) Animal localization of a transcript encoding a signaling molecule required for axis development in the egg of a cnidarian, *Clytia*. (**C**) mRNA enrichment in synapses of an *Aplysia* sensory neuron in response to contact with a target motor neuron (blue). (**D**) Apical localization of an mRNA in the *Drosophila* embryo, which facilitates entry of its transcription factor product into the nuclei (purple). (**E**) mRNA localization in pseudopodial protrusions of a cultured mammalian fibroblast (red signal indicates the cell volume). (**F**) mRNA enrichment within a *Xenopus* axonal growth cone. mRNAs were visualized by means of in situ hybridization except in (E), in which the MS2–green fluorescent protein (GFP) system was used. *Drosophila* images are reproduced from (*50*) with permission. [Images were kindly provided by (B) T. Momose and E. Houliston, (C) D.O. Wang and K. Martin, (D) M. Dienstbier, (E) S. Mili and I. Macara, and (F) F. van Horck.]

Polarized Functions in Other Cell Types

The functional consequences of disrupting mRNA localization have now been tested in many other cell types. These studies have revealed an important role for the localiza-

tion of specific mRNAs in facilitating subcellular protein localization, helping to establish or maintain cell polarity (table S2).

A particularly intriguing example comes from primary chicken fibroblasts. Here, interference with the β-actin zip code through antisense oligonucleotides strongly reduces the persistence of cell movement (*46*). But given that the number of protein molecules encoded by the localizing mRNAs represents only a tiny fraction of the total β-*actin* protein near the leading edge, why is mRNA targeting important? It is conceivable that newly synthesized β-actin monomers polymerize more efficiently than older copies, for instance, because of differential posttranslational modifications or modulation by chaperones. An alternative explanation relates to the potential for transport along a cytoskeletal track to convey multiple β-*actin* mRNA molecules to a small region of the cytoplasm. This could dictate a high local concentration of the protein, aiding rapid polymerization of filaments. The finding that all seven transcripts encoding Arp2 and Arp3 components are localized behind the leading edge lends support to the notion that mRNA targeting controls actin dynamics by facilitating the local assembly of protein complexes (*47*).

But it is not just mRNAs encoding cytoskeletal proteins that are localized in dynamic cells. At least 50 transcript species, coding for proteins with diverse functions, are enriched in pseudopodial protrusions of mouse fibroblasts in response to migratory stimuli (*48*). The localization mechanism is microtubule-associated and appears to be distinct from that used to target mRNAs behind the leading edge of chicken fibroblasts, involving direct roles of the adenomatous polyposis coli (APC) tumor suppressor and FMRP. This study, together with others, opens up new perspectives for elucidating links between mRNA localization and human disease [supporting online material (SOM) text].

Perspectives

Key principles of mRNA localization mechanisms in animal cells have now been established and many players identified. An important challenge is to piece together a detailed molecular understanding of the interactions that govern the recognition and differential sorting of mRNAs as well as the interplay with translational regulators. In cases in which mRNA localization is regulated by extrinsic cues (Fig. 3), what aspects of the translocation process are being targeted and how? And what is the copy number of mRNAs within the majority of mRNPs (SOM text)? Addressing these questions will benefit from insights from genetically tractable model organisms, including flies and fungi, and from advances in the ability to visualize the composition and behavior of mRNPs in living cells. The use of unbiased genome-wide methods to identify binding sites for specific transacting factors (*49*) could also have profound effects on our understanding of the recognition of localizing mRNAs.

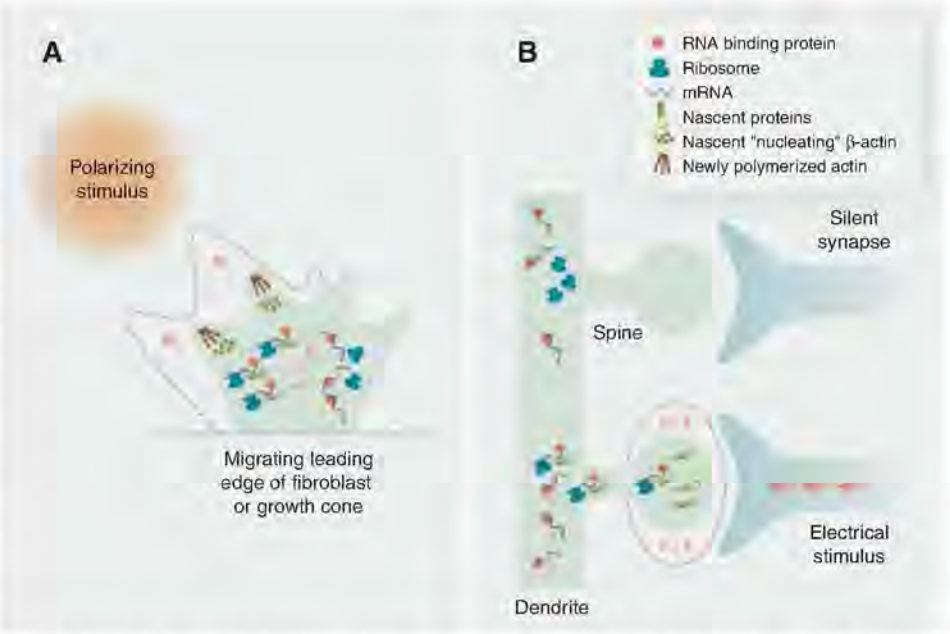


Fig. 3. Extrinsic stimuli elicit changes in subcellular mRNA localization and translation. **(A)** A polarizing stimulus elicits asymmetric localization and translation of mRNAs encoding β-actin and actin regulators on the near-stimulus side of the leading edge of migrating cells, such as fibroblasts and axonal growth cones, thus contributing to polarized cell movement and directional steering. The dashed outline denotes the post-stimulus trajectory. **(B)** Electrical input from presynaptic contacts selectively induces localized trafficking and translation of specific mRNAs in dendrites that mediate changes in spine morphology (dashed outline) and plasticity. Several aspects of these models are speculative.

A large number of studies have now highlighted the importance of subcellular mRNA localization in diverse cellular processes. Nonetheless, several questions remain from a functional perspective. What are the relative contributions of mRNA localization and localized translation to processes such as axon guidance, synaptogenesis, and synaptic plasticity? What is the extent and importance of asymmetric targeting of microRNAs? The requirements for some localizing mRNAs are independent from their translation (table S2); could this reflect a widespread structural role for mRNA in facilitating the assembly of protein complexes?

References and Notes

1. W. R. Jeffery, C. R. Tomlinson, R. D. Brodeur, *Dev. Biol.* **99**, 408 (1983).
2. J. B. Lawrence, R. H. Singer, *Cell* **45**, 407 (1986).
3. E. Lecuyer *et al.*, *Cell* **131**, 174 (2007).
4. K. Zarnack, M. Feldbrugge, *Mol. Genet. Genomics* **278**, 347 (2007).
5. T. W. Okita, S. B. Choi, *Curr. Opin. Plant Biol.* **5**, 553 (2002).
6. K. M. Forrest, E. R. Gavis, *Curr. Biol.* **13**, 1159 (2003).
7. A. Bashirullah, R. L. Cooperstock, H. D. Lipshitz, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7025 (2001).
8. U. Wolke, G. Weidinger, M. Koprunner, E. Raz, *Curr. Biol.* **12**, 289 (2002).
9. J. Y. Hu, X. Meng, S. Schacher, *J. Neurosci.* **22**, 2669 (2002).
10. J. H. Carson, H. Cui, E. Barbarese, *Curr. Opin. Neurobiol.* **11**, 558 (2001).
11. S. L. Bullock, A. Nicol, S. P. Gross, D. Zicha, *Curr. Biol.* **16**, 1447 (2006).
12. J. L. Dynes, O. Steward, *J. Comp. Neurol.* **500**, 433 (2007).
13. V. L. Zimyanin *et al.*, *Cell* **134**, 843 (2008).
14. T. J. Messitt *et al.*, *Dev. Cell* **15**, 426 (2008).
15. J. Krauss, S. Lopez de Quinto, C. Nusslein-Volhard, A. Ephrussi, *Curr. Biol.* **19**, 1058 (2009).
16. V. Van De Bor, E. Hartswood, C. Jones, D. Finnegan, I. Davis, *Dev. Cell* **9**, 51 (2005).
17. N. Vogt, I. Koch, H. Schwarz, F. Schnorrer, C. Nusslein-Volhard, *Development* **133**, 3963 (2006).
18. K. Babu, Y. Cai, S. Bahri, X. Yang, W. Chia, *Genes Dev.* **18**, 138 (2004).
19. J. K. Yisraeli, S. Sokol, D. A. Melton, *Development* **108**, 289 (1990).
20. R. Delanoue, I. Davis, *Cell* **122**, 97 (2005).
21. T. T. Weil, K. M. Forrest, E. R. Gavis, *Dev. Cell* **11**, 251 (2006).
22. D. Fusco *et al.*, *Curr. Biol.* **13**, 161 (2003).
23. L. Chang, Y. Shav-Tal, T. Trcek, R. H. Singer, R. D. Goldman, *J. Cell Biol.* **172**, 747 (2006).
24. A. Jambhekar, J. L. Derisi, *RNA* **13**, 625 (2007).
25. M. Dienstbier, F. Boehl, X. Li, S. L. Bullock, *Genes Dev.* **23**, 1546 (2009).
26. P. Chartrand, X. H. Meng, S. Huttelmaier, D. Donato, R. H. Singer, *Mol. Cell* **10**, 1319 (2002).
27. F. Besse, A. Ephrussi, *Nat. Rev. Mol. Cell Biol.* **9**, 971 (2008).
28. S. Huttelmaier *et al.*, *Nature* **438**, 512 (2005).
29. F. Pan, S. Huttelmaier, R. H. Singer, W. Gu, *Mol. Cell. Biol.* **27**, 8340 (2007).
30. R. A. Lewis, J. A. Gagnon, K. L. Mowry, *Mol. Cell. Biol.* **28**, 678 (2008).
31. C. Giorgi, M. J. Moore, *Semin. Cell Dev. Biol.* **18**, 186 (2007).
32. C. Giorgi *et al.*, *Cell* **130**, 179 (2007).
33. M. L. King, T. J. Messitt, K. L. Mowry, *Biol. Cell* **97**, 19 (2005).
34. C. Sardet, P. Dru, F. Prodon, *Biol. Cell* **97**, 35 (2005).
35. A. Amiel, E. Houliston, *Dev. Biol.* **327**, 191 (2009).
36. A. Jedrusik *et al.*, *Genes Dev.* **22**, 2692 (2008).
37. J. D. Lambert, L. M. Nagy, *Nature* **420**, 682 (2002).
38. O. Steward, C. S. Wallace, G. L. Lyford, P. F. Worley, *Neuron* **21**, 741 (1998).

39. C. R. Bramham, P. F. Worley, M. J. Moore, J. F. Guzowski, *J. Neurosci.* **28**, 11760 (2008).
40. S. Miller *et al.*, *Neuron* **36**, 507 (2002).
41. V. Lyles, Y. Zhao, K. C. Martin, *Neuron* **49**, 349 (2006).
42. D. O. Wang *et al.*, *Science* **324**, 1536 (2009).
43. K. M. Leung *et al.*, *Nat. Neurosci.* **9**, 1247 (2006).
44. J. Yao, Y. Sasaki, Z. Wen, G. J. Bassell, J. Q. Zheng, *Nat. Neurosci.* **9**, 1265 (2006).
45. L. J. Cox, U. Hengst, N. G. Gurskaya, K. A. Lukyanov, S. R. Jaffrey, *Nat. Cell Biol.* **10**, 149 (2008).
46. E. H. Kislauskis, X. Zhu, R. H. Singer, *J. Cell Biol.* **136**, 1263 (1997).
47. L. A. Mingle *et al.*, *J. Cell Sci.* **118**, 2425 (2005).
48. S. Mili, K. Moissoglu, I. G. Macara, *Nature* **453**, 115 (2008).
49. D. D. Licatalosi *et al.*, *Nature* **456**, 464 (2008).
50. S. L. Bullock, *Semin. Cell Dev. Biol.* **18**, 194 (2007).
51. We apologize to those whose primary work could not be cited because of space constraints. We acknowledge the continued importance of research in yeast for informing and inspiring studies on mRNA localization in animal cells. We thank C. Dix, A. Lin, and F. van Horck for comments on the manuscript, those who provided images, and many colleagues for answering queries. Work in our laboratories is supported by a Wellcome Trust Programme grant (C.H.) and the MRC and a Lister Institute Prize Fellowship (S.B.).

10.1126/science.1176488

REVIEW

The Extracellular Matrix: Not Just Pretty Fibrils

Richard O. Hynes

The extracellular matrix (ECM) and ECM proteins are important in phenomena as diverse as developmental patterning, stem cell niches, cancer, and genetic diseases. The ECM has many effects beyond providing structural support. ECM proteins typically include multiple, independently folded domains whose sequences and arrangement are highly conserved. Some of these domains bind adhesion receptors such as integrins that mediate cell-matrix adhesion and also transduce signals into cells. However, ECM proteins also bind soluble growth factors and regulate their distribution, activation, and presentation to cells. As organized, solid-phase ligands, ECM proteins can integrate complex, multivalent signals to cells in a spatially patterned and regulated fashion. These properties need to be incorporated into considerations of the functions of the ECM.

All cells make close contact with the extracellular matrix (ECM), either continuously or at important phases of their lives (for instance, as stem or progenitor cells or during cell migration and invasion). The ECM is well known for its ability to provide structural support for organs and tissues, for cell layers in the form of basement membranes, and for individual cells as substrates for migration. The role of the ECM in cell adhesion and signaling to cells through adhesion receptors such as integrins has received much attention (*1*–*3*), and, more recently, mechanical characteristics of the matrix (stiffness, deformability) have also been recognized to provide inputs into cell behavior (*4*, *5*). Thus, ECM proteins and structures play vital roles in the determination, differentiation, proliferation, survival, polarity, and migration of cells. ECM signals are arguably at least as important as soluble signals in governing these processes, and probably more so. Here, I will emphasize different contributions of the ECM and ECM proteins to cell and tissue behavior, namely their roles in binding, integrating, and presenting growth factor signals to cells.

The Complex Domain Structures of ECM Proteins

There are hundreds of ECM proteins encoded in vertebrate genomes. Many of the genes are ancient, such as those composing the basement membrane toolkit (type IV collagens, laminins, nidogen, perlecan, and type XV/XVIII collagen), which is found in most metazoa, and one can argue that basement membranes were crucial to the evolution of metazoa (*6*). However, many vertebrate ECM proteins and genes evolved much more recently, during evolution of the deuterostome lineage, and that expansion includes not only elaboration of preexisting families (for example, laminins and collagens) but also novel proteins [e.g., fibronectins (FNs) and tenascins]. What purposes are served by this proliferation of ECM proteins? ECM proteins are large and complex, with multiple distinct domains, and are highly conserved among different taxa (Fig. 1). It is not necessary for proteins to be large or complex to generate strong, stable fibrils—intermediate filament proteins and type I collagen provide notable examples to the contrary. So why are most ECM proteins so large, complex, and conserved? Many ECM proteins have dozens of individually folded domains, but in most cases, we understand the functions of only a few of them. What is the purpose of the other domains? The conserved domains are arranged in specific juxtapositions, sometimes controlled by highly regulated alternative splicing. The clear implication is that the specific domains and architectures of ECM proteins contain information of biological importance and evolutionary value. This article will explore that hypothesis in light of recent discoveries concerning representative ECM proteins.

Howard Hughes Medical Institute, David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: rohynes@mit.edu

ECM Proteins and Growth Factor Signaling

One long-standing idea is that the ECM binds growth factors, which is certainly true. Many growth factors [e.g., fibroblast growth factors (FGFs) and vascular endothelial growth factors (VEGFs)] bind avidly to heparin and to heparan sulfate, a component of many ECM proteoglycans (PGs). Hence, a generally held view is that heparan sulfate PGs act as a sink or reservoir of growth factors and may assist in establishing stable gradients of growth factors bound to the ECM; such gradients of morphogens play vital roles in patterning developmental processes. It is also often proposed that growth factors can be released from the ECM by degradation of ECM proteins or of the glycosaminoglycan components of PGs. Those models place the ECM in a distal role, acting as localized reservoirs for soluble growth factors that will be released from the solid phase to function as traditional, soluble ligands. However, some growth factors actually bind to their signaling receptors with heparan sulfate as a cofactor. The binding of FGF to its receptor (FGFR) depends on a heparan sulfate chain binding simultaneously (*7*), and transforming growth factor–β (TGF-β) ligands bind first to integral-membrane PGs that "present" these ligands during signaling (*8*); effectively they act as solid-phase ligands. Such phenomena may well be more widespread than the few, well-studied examples that are currently known. There are also increasing numbers of examples of growth factors binding to ECM proteins themselves, without the involvement of glycosaminoglycans, supporting the notion that the presentation of growth factor signals by ECM proteins is an important part of ECM function.

There are several related concepts that need to be kept separate in thinking about and analyzing functions of the ECM in signaling to cells. First, standard ECM receptors, such as integrins and discoidin domain tyrosine kinase receptors, are themselves signal transduction receptors. Their ligands are specific domains and motifs embedded in the ECM proteins, and ECM-integrin interactions lead to signal transduction responses that are at least as complex and important as those triggered by soluble ligands such as EGF, platelet-derived growth factor, and VEGF (*1*–*3*). Second, and less clearly, there are numerous reports of "cross talk"

and "synergy" between signaling by integrins and by various growth factors (*9*). In most cases, it is uncertain whether such cross talk involves (i) membrane-proximal interactions or (ii) cooperation in the downstream signal transduction pathways. Another concept is that intrinsic domains within ECM proteins might act as ligands for canonical growth factor receptors. This suggestion arose from the observation that laminin contains multiple EGF-like domains, as do many ECM proteins [e.g., laminins, tenascins, thrombospondins (TSPs), fibrillins; see Fig. 1], which might bind to EGF receptors and signal as solid-phase ligands (*10*). EGF-like domains from laminin (*11*, *12*) or tenascin (*13*) presented as soluble ligands can bind to EGFR and modulate its signaling, and it is often hypothesized that fragments of ECM proteins can be released by proteolysis (for instance, by matrix metalloproteases) and act as soluble ligands, similar to the idea that matrix-bound growth factors can be released by ECM degradation. In both cases, the ECM acts as a reservoir of growth factors (bound or intrinsic), which can be released as soluble factors to bind their receptors. However, the interesting idea that intrinsic growth factor–like ligands can act from the solid-phase deserves more intensive investigation and careful experimental distinction from alternatives such as release of bound or intrinsic ligands. We will explore this idea and the related concept that ECM proteins bind and present growth factors as organized solid-phase ligands.

Growth Factor Binding to ECM Proteins

There is increasing evidence for specific, direct binding of growth factors to ECM proteins (*13*, *14*). Both FN and vitronectin bind hepatocyte growth factor (HGF) and form complexes of Met (the HGF receptor) and integrins (the ECM receptors), leading to enhanced cell migration (*15*). Similarly VEGF binds to specific FN type III (FN3) domains in both FN and tenascin-C, and these associations promote cell proliferation (*16*, *17*). In the case of the FN-VEGF binding, the effect on proliferation requires the binding sites for integrins and VEGF to be in the same molecule, suggesting a requirement for juxtaposition of the two receptors (integrin α5β1 and VEGFR2), rather than merely downstream cross talk (*16*). FN3 domains are prevalent in many ECM proteins, and membrane receptors and their potential for binding soluble factors need further investigation.

Other widely distributed ECM domains can bind and present growth factors. *Drosophila* collagen IV binds Dpp [a bone morphogenetic protein (BMP) homolog] and enhances its interactions with BMP receptors; this collagen-BMP interaction is crucial in regulating the dorsoventral axis and the numbers of germinal stem cells in the ovary, both processes that are dependent on gradients of Dpp (*18*). Collagen IV is a universal constituent of basement membranes, and the key Dpp-binding motif identified in the C-terminal domains of the two *Drosophila* collagen IV subunits is highly conserved across phyla, suggesting that this interaction may be important in other contexts (*18*). Another instructive example is collagen II, the major collagen of cartilage that, near its N terminus, contains a chordin-like VWC domain that binds TGF-β1 and BMP-2, two chondrogenic growth factors. The VWC domain is alternatively spliced, included in prechondrogenic mesoderm and early developing cartilage but excluded in mature cartilage (*19*). The VWC or chordin domain is found in many ECM proteins and in known regulators of BMPs, and it typically acts as a negative regulator of their functions (*20*). These examples illustrate the capacity of conserved elements of ECM proteins to regulate, either positively or negatively, the functions of diffusible morphogens of the BMP family.

TGF-β Regulation by ECM Binding

The regulation of TGF-β signaling by ECM proteins is one of the best developed examples of this capacity. Each of the precursors of TGF-β isoforms 1 to 3 is cleaved by a furin protease to the mature TGF-β and its propeptide, known as latency-associated peptide (LAP). The LAP and TGF-β remain noncovalently associated in a complex called the small latency complex (SLC), and in this form, TGF-βs are inactive (*21*, *22*). The LAPs then S-S bond to one of the latent TGF-β–binding proteins (LTBPs) to form large latent complexes (LLCs), and many cells secrete TGF-βs already assembled into such complexes. In turn, the LTBPs bind to other ECM proteins (including fibrillins and FNs), thereby incorporating the different TGF-β isoforms into extracellular matrices in latent form (Figs. 1 and 2A). LTBP-mediated incorporation into the ECM is necessary for subsequent effective activation of TGF-βs. There are several mechanisms for activation (Fig. 2B), including degradation of ECM proteins such as fibrillin or LTBPs. Activation can also occur by cleavage or conformational change in LAP, exposing or releasing the TGF-βs to bind and activate their receptors (*21*, *22*). Another ECM protein, TSP, can activate TGF-βs by binding and dissociating LAP or by activating metallopro-

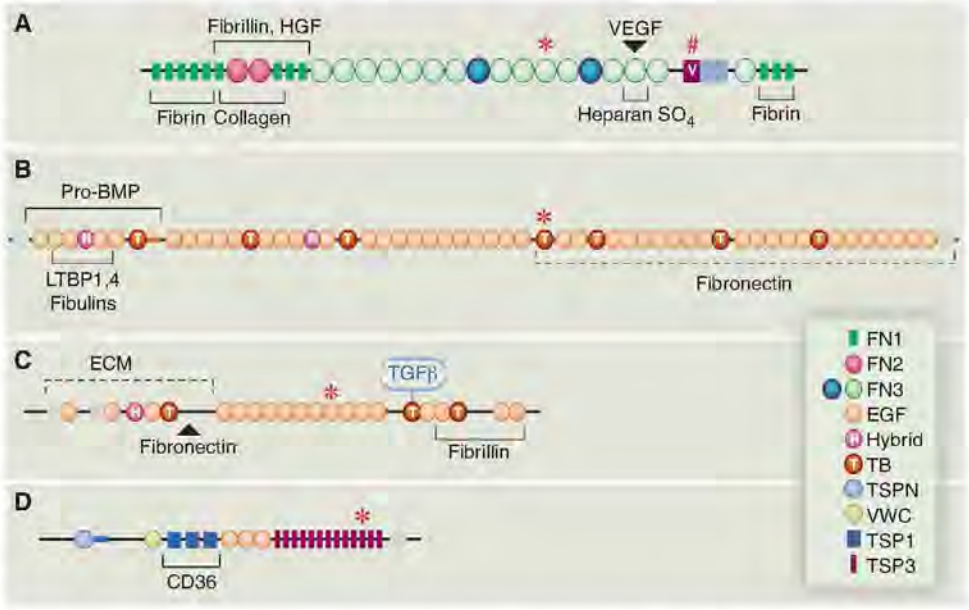


Fig. 1. The complex domain structures of ECM proteins. Representative ECM proteins illustrating multiple, independently folded domains, which occur in differing combinations in different ECM proteins through exon shuffling during evolution. Domain structures were generated with SMART (*36*) and edited for details of individual proteins. (**A**) Fibronectin. Encoded by a single gene but alternatively spliced at three regions [blue circles and box and V (variable) segment] to generate 12 proteins in rodents and 20 in humans. FN3 domains are widespread in ECM proteins. Binding sites for other matrix proteins are marked. The heparan sulfate–binding site can interact with PGs or with syndecan, an integral-membrane PG. Integrin-binding sites; RGD (indicated by an asterisk) and LDV (Leu-Asp-Val, indicated by a pound sign). FN is a proangiogenic molecule, whose function depends on both the RGD site and the two alternatively spliced FN3 domains (*37*, *38*). FN also binds the proangiogenic growth factors VEGF and HGF (*16*, *17*). (**B**) Fibrillin-1. Fibrillins include EGF-like domains, found in many ECM proteins, as well as TB (TGFβ-binding, denoted by T) and hybrid (H) domains, specific to fibrillins and LTBPs (*21*, *22*). Binding sites for other matrix proteins and growth factors are marked. (**C**) LTBP-1. Four-gene family with structures related to fibrillins. Known binding sites for TGF-β/LAP latent complex (SLC, blue), fibrillin, and FN are marked. RGD (asterisk) sequences in fibrillins and LTBPs may bind integrins. (**D**) Thrombospondin-1 (TSP-1). TSPs contain TSP1 repeats (also found in other ECM proteins), EGF-like repeats, and a VWC domain, known in other proteins to bind BMPs. TSP3 repeats (purple) and C-terminal domains are unique to TSPs and bind multiple Ca^{++} ions. The RGD (asterisk) sequence is known to bind to integrins. TSPs 1 and 2 have the structure shown, and both have antiangiogenic activity located in the TSP1 repeats, which bind to the CD36 receptor (*39*).

teases; mice lacking TSP-1 develop pneumonia because of reduced levels of active TGF-β in their lungs (*23*). Yet another mechanism for activation of TGF-βs involves αvβ6 and αvβ8 integrins, which bind to Arg-Gly-Asp (RGD) sequences in LAP1 and LAP3 (*24*, *25*). αvβ8 integrin appears to cooperate with metalloproteases to release TGF-β. However, αvβ6 integrin activates TGF-β without any requirement for proteolysis. Instead, it binds to LAP and, in the presence of mechanical strain between the cells expressing the integrin and the ECM to which the SLC is attached, deforms LAP to expose the associated TGF-β (Fig. 2B). The activated TGF-β is not released in soluble, diffusible form but appears to act only at short range, perhaps as a bound solid-phase ligand. Thus, the binding, sequestration in latent form, and subsequent activation of TGF-βs all intimately involve a variety of ECM proteins (Fig. 2). The whole assemblage acts like a regulated machine incorporating both negative and positive regulation; incorporation of TGF-β into the matrix anchors and localizes the growth factor in a latent form, which can subsequently be locally activated by proteolysis or by mechanical strain (*21*–*25*). Mutations in many of the ECM proteins, integrins, and the RGD sites in the LAPs confirm the relevance of these interactions in vivo.

Further analyses of LAPs, LTBPs, and fibrillins show that the TGF-β–LAP complex binds to LTBP-1 through a specific TGF-binding (TB) domain and adjacent EGF domains (Fig. 1). TB domains, as well as hybrid domains (hybrids of TB and EGF domains), are unique to fibrillins and LTBPs, and there are several in each of those proteins, suggesting that they may be able to bind other BMP family members (Fig. 1). Indeed, proBMP-7 can bind to fibrillin-1 in an N-terminal region containing a hybrid and a TB domain (*26*). Furthermore, fibrillin-2 and BMP-7 mutations interact in causing syndactyly and polydactyly in mice (*27*), and a related human disease, congenital contractual arachnodactyly, arises from mutations in fibrillin-2 (*28*, *29*). Other functionally important interactions between members of the TGF/BMP and LTBP/fibrillin families probably remain to be discovered. The interactions of different LTBPs and fibrillins with diverse TGF/BMP family members potentially target different signals to different locations.

The implications of ECM-based regulation of TGF-β function for human disease have recently become abundantly clear in the case of Marfan syndrome, a genetic disease resulting from mutations in the gene for fibrillin-1 (*28*, *29*). Like many other genetic diseases whose target genes encode ECM proteins, this disease is associated with defective assembly of ECM components—in this case, the microfibrils of which fibrillins are components. The phenotype was originally attributed to mechanical consequences of these structural defects. However, the known associations of fibrillins with LTBPs suggested that activation of TGF-βs might also play a role. In mouse models of Marfan syndrome, activation of TGF-β is markedly increased, and many of the phenotypic consequences of mutations in fibrillin-1 can be ameliorated by TGF-β antagonists, an insight that already has clinical applications (*28*, *29*).

ECM Proteins as Localized, Multivalent Signal Integrators

Thus, discrete domains in ECM proteins can bind and regulate functions of canonical growth factors. Many such domains are found in multiple ECM proteins in different combinations and arrangements, and presumably, many more ECM/growth factor interactions remain to be discovered. Other domains and motifs in these ECM proteins have the potential to bind directly to cell surface–adhesion receptors such as integrins. At the very least, the coexistence in the same ECM proteins of sites for cell adhesion and binding sites for growth factors concentrates the growth factors close to their own cell surface receptors. Thus, localization of growth factors at the cellular level by binding to the ECM can localize their signaling, and binding of growth factors to the ECM probably contributes to establishment of stable gradients. According to this model, morphogen gradients are composed jointly of soluble, diffusible factors and the ECM—and both are necessary. ECM-bound growth factors could be released locally or presented as complexes still bound to the ECM proteins; as mentioned earlier, there is also the potential (as yet unproven) for specific intrinsic domains in ECM proteins (such as EGF-like domains) to bind directly to growth factor receptors.

ECM proteins are highly conserved, not only in the sequences of specific domains but also in the arrangements of those domains. Furthermore, specific domains are often inserted or omitted by regulated alternative splicing, thus changing the complement of domains. This could alter the binding of specific growth factors, as in the case of the VWC domain in type II collagen (*19*), or interactions with cell surface receptors. In the case of agrin, inclusion of two small exons confers on agrin the ability to bind to heparan sulfate and dystroglycan and greatly enhances the clustering of acetylcholine receptors (*30*). ECM proteins can also synergize with growth factors in affecting cell proliferation and migration (*9*). Although such synergy does not in principle require juxtaposition, experiments on VEGF binding by FN show that the synergy requires the binding sites for integrins and VEGF to be coupled in the same molecule—presenting them as two separate, substrate-bound fragments of FN does not suffice (*16*). If such proximity is important, ECM molecules, by virtue of their ordered-domain organization, could act to organize complexes of receptors in the plane of the membrane. Such complexes could enhance membrane-proximal regulation among the receptors and promote integration of the signals transduced (Fig. 3). An instructive parallel can be found in the clustering of immunoregulatory receptors in immunological synapses [which also involve cross talk among integrins and other receptors (*31*, *32*)]. Immunological synapses have substructure: Different receptors occupy different zones within the synapse. ECM-mediated clusters could have highly detailed substructure, and the juxtaposition of different receptors could be driven by the arrangement of domains in the ECM protein at a resolution of several nanometers. One could think of ECM proteins and their associated partners (growth factors and other ECM proteins) as solid-phase growth factors metaphorically playing chords, in contrast with soluble growth factors that one could view as playing single notes (Fig. 3).

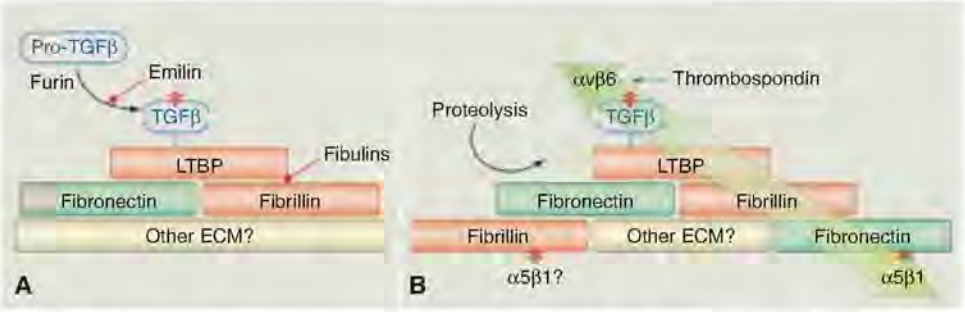


Fig. 2. ECM interactions regulating TGF-β. (**A**) Incorporation into the ECM. Cleavage by furin protease of Pro–TGF-β to the small latent complex (SLC) comprising TGF-β and LAP (blue) is inhibited by emilin, an ECM protein. The SLC binds to LTBP, via S-S bonding to a TB domain, to form the LLC, in which form the TGF-β is inactive (*21*, *22*). LTBP then binds to fibrillin and to FN (see Fig. 1 for specific interaction domains). Fibulins compete for LTBP binding to fibrillin (*40*). Fibrillin binds to preexisting FN fibrils or assembles into microfibrils, and both fibrillin and FN undergo further homomeric and heteromeric interactions within the ECM. (**B**) Activation of ECM-bound latent TGF-β. TGF-β can be activated by proteolysis of the ECM proteins and/or of LAP or directly by thrombospondin (see text). TGF-β can also be activated by mechanical strain (large green arrow). This strain arises from cytoskeletal force applied through αvβ6 integrin, which binds to an RGD site in LAP and requires attachment of the TGF-β/LAP complex through LTBP to the FN-rich matrix, which, in turn, is attached via α5β1 integrin to other cells. Fibrillin might also be attached to cells via integrins.

The very nature of the ECM imposes spatial context on the signaling. Cells are often polarized by their associations with the ECM—the basement membranes to which epithelial sheets attach define the base and polarity of the cells and confer ability to respond to soluble growth factors such as EGF. The deformability of the ECM also affects the responses of cells (*2*–*4*, *33*, *34*). ECM molecules are flexible and extendable, and mechanical tension can uncover cryptic sites within them (*35*). Such mechanically exposed cryptic sites could bind additional cell surface receptors or growth factors. Mechanical extension or the inclusion or exclusion of alternatively spliced domains could also alter the physical relations among other domains, thus affecting the composition and spatial arrangement of the hypothesized organized patches of receptors.

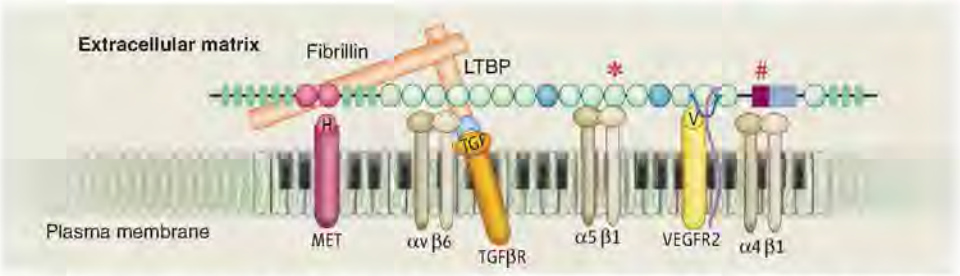


Fig. 3. Multidomain interactions of ECM proteins with cells. The example shown is FN (*41*). Multiple domains are known to bind to integrins, other ECM proteins, and growth factors, as shown. Integrins α5β1 and α4β1 bind, respectively, to RGD and LDV motifs; heparan sulfate chains of syndecan (purple/blue) bind to FN3-13 as does VEGF. Evidence suggests that VEGF (V, yellow) signals through its own receptor (VEGFR2) more effectively when bound to FN (*16*). The same is proposed here for HGF (H) and its receptor (Met, pink). As shown in Figs. 1 and 2, fibrillin binds to an N-terminal region of FN and, in turn, binds LTBP, which recruits TGF-β in a latent complex with LAP (blue crescent). αvβ6 integrin can bind an RGD site in LAP, activating TGF-β, so that it can bind its own receptors (orange). The proposal is that FN organizes and integrates all these signals at two levels. First, by recruiting growth factors to the ECM, FN localizes those signals at the cellular level. Second, the close juxtaposition of the domains in FN brings the different receptors together into an organized submicron patch in the cell surface membrane. Each domain is 2 to 4 nm in diameter, and the entire FN subunit shown is 60 to 70 nm long, so the receptors will be brought into close apposition such that their signals provide complex, integrated information to the cell—metaphorically generating "chords" and "melodies" in contrast with the "single notes" generated by each receptor. FN is essential for angiogenesis, and most of the bound receptors and ligands have been shown to play roles in angiogenesis. This model suggests that FN and its associated ECM proteins orchestrate and integrate these signals. In addition, alternatively spliced domains of FN (blue circles; see Fig. 1A) are also necessary for proper vascular development, and it is a reasonable hypothesis that they introduce additional ligands and/or receptors into the mix.

Implications for Future Research

The ideas explored here need further experimental tests. There are relatively few well-documented examples of specific growth factor binding by domains in ECM proteins, but this possibility could be readily investigated. There are even fewer cases where it is clear whether ECM-bound growth factors need to be released to soluble form or can act as solid-phase ligands. The proposition that intrinsic domains of ECM proteins can directly affect canonical growth factor receptors, either as solid-phase ligands or as locally released soluble ligands, needs more study. The idea that specific arrangements of domains confer important information can be tested. The possible effects of mechanical strain on exposure of cryptic binding sites for growth factors, receptors, or other ECM proteins are just beginning to be explored. The nature of ECM-induced receptor complexes in the membrane can be investigated by methods such as single-molecule tracking, fluorescence energy transfer methods, correlation microscopy, high-resolution electron microscopy, and chemical cross-linking. The effects of regulated alternative splicing of ECM proteins on all of these questions and the implications of the diversity within families of proteins such as LTBPs and fibrillins need to be investigated further.

The ECM is a fundamental component of the microenvironment of cells and has been substantially expanded during the evolution of vertebrates. Some of that elaboration has contributed to structural components such as bones and teeth, but it is evident that this is only one role of the ECM. The ECM provides much more than mechanical support and a locus for cell adhesion, with potential roles in basement membranes, stem cell niches, and tumors. All epithelial cells are in association with basement membranes for at least part of their lives, and many stem cell niches include the ECM. ECM composition and organization undergo radical alterations in cancer and could affect survival, proliferation, and other properties of both tumor and stromal cells. Ever since McKusick's initial cataloguing of a diverse set of genetic diseases affecting the ECM more than 50 years ago, it has been implicitly assumed that the pathological consequences were a direct result of defects in ECM assembly. Although those defects do exist, and no doubt contribute, in Marfan syndrome and related diseases, many phenotypic consequences are indirect effects of dysregulation of TGF-β signaling consequent on the ECM defects. Structural defects are difficult to treat in the absence of gene therapy or stem cell therapies, but growth factor signaling offers simpler and more accessible targets for intervention. Further investigations of the roles of ECM proteins in regulating signaling events should yield additional leads.

References and Notes

1. R. O. Hynes, *Cell* **110**, 673 (2002).
2. A. L. Berrier, K. M. Yamada, *J. Cell. Physiol.* **213**, 565 (2007).
3. K. R. Legate, S. A. Wickstrom, R. Fassler, *Genes Dev.* **23**, 397 (2009).
4. D. E. Discher, D. J. Mooney, P. W. Zandstra, *Science* **324**, 1673 (2009).
5. B. Geiger, J. P. Spatz, A. D. Bershadsky, *Nat. Rev. Mol. Cell Biol.* **10**, 21 (2009).
6. C. A. Whittaker *et al.*, *Dev. Biol.* **300**, 252 (2006).
7. M. Mohammadi, S. K. Olsen, R. Goetz, *Curr. Opin. Struct. Biol.* **15**, 506 (2005).
8. Y. Shi, J. Massague, *Cell* **113**, 685 (2003).
9. N. Alam *et al.*, *J. Cell. Physiol.* **213**, 649 (2007).
10. J. Engel, *FEBS Lett.* **251**, 1 (1989).
11. G. Panayotou, P. End, M. Aumailley, R. Timpl, J. Engel, *Cell* **56**, 93 (1989).
12. S. Schenk *et al.*, *J. Cell Biol.* **161**, 197 (2003).
13. A. K. Iyer, K. T. Tran, L. Griffith, A. Wells, *J. Cell. Physiol.* **214**, 504 (2007).
14. G. G. Vaday, O. Lider, *J. Leukoc. Biol.* **67**, 149 (2000).
15. S. Rahman *et al.*, *BMC Cell Biol.* **6**, 8 (2005).
16. E. S. Wijelath *et al.*, *Circ. Res.* **99**, 853 (2006).
17. T. Ishitsuka, T. Ikuta, H. Ariga, K. I. Matsumoto, *Biol. Pharm. Bull.* **32**, 1004 (2009).
18. X. Wang, R. E. Harris, L. J. Bayston, H. L. Ashe, *Nature* **455**, 72 (2008).
19. Y. Zhu, A. Oganesian, D. R. Keene, L. J. Sandell, *J. Cell Biol.* **144**, 1069 (1999).
20. J. Garcia Abreu *et al.*, *Gene* **287**, 39 (2002).
21. D. B. Rifkin, *J. Biol. Chem.* **280**, 7409 (2005).
22. P. ten Dijke, H. M. Arthur, *Nat. Rev. Mol. Cell Biol.* **8**, 857 (2007).
23. J. Lawler *et al.*, *J. Clin. Invest.* **101**, 982 (1998).
24. D. Sheppard, *Cancer Metastasis Rev.* **24**, 395 (2005).
25. P. J. Wipff, B. Hinz, *Eur. J. Cell Biol.* **87**, 601 (2008).
26. K. E. Gregory *et al.*, *J. Biol. Chem.* **280**, 27970 (2005).
27. E. Arteaga-Solis *et al.*, *J. Cell Biol.* **154**, 275 (2001).
28. F. Ramirez, L. Y. Sakai, D. B. Rifkin, H. C. Dietz, *Cell. Mol. Life Sci.* **64**, 2437 (2007).
29. F. Ramirez, H. C. Dietz, *J. Biol. Chem.* **284**, 14677 (2009).
30. J. T. Campanelli, G. G. Gayer, R. H. Scheller, *Development* **122**, 1663 (1996).
31. A. Kupfer, H. Kupfer, *Semin. Immunol.* **15**, 295 (2003).
32. M. L. Dustin, *Annu. Rev. Cell Dev. Biol.* **24**, 577 (2008).
33. A. J. Engler, S. Sen, H. L. Sweeney, D. E. Discher, *Cell* **126**, 677 (2006).
34. V. Vogel, M. P. Sheetz, *Curr. Opin. Cell Biol.* **21**, 38 (2009).
35. V. Vogel, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 459 (2006).
36. SMART database (http://smart.embl-heidelberg.de/).
37. S. Takahashi *et al.*, *J. Cell Biol.* **178**, 167 (2007).
38. S. Astrof, D. Crowley, R. O. Hynes, *Dev. Biol.* **311**, 11 (2007).
39. J. C. Adams, J. Lawler, *Int. J. Biochem. Cell Biol.* **36**, 961 (2004).
40. R. N. Ono *et al.*, *J. Biol. Chem.* **284**, 16872 (2009).
41. R. O. Hynes, *Fibronectins* (Springer, New York, 1990).

42. I thank A. Naba and K. Certel for constructive criticisms of the text and figures, and I gratefully acknowledge support from the Howard Hughes Medical Institute and NIH.

10.1126/science.1176009

REVIEW

Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart

John D. Scott[1]* and Tony Pawson[2,3]*

Signal transduction can be defined as the coordinated relay of messages derived from extracellular cues to intracellular effectors. More simply put, information received on the cell surface is processed across the plasma membrane and transmitted to intracellular targets. This requires that the activators, effectors, enzymes, and substrates that respond to cellular signals come together when they need to.

The quest to discover the full complement of cell-signaling components has achieved notable success, and so the next challenge is to establish how these pieces work in concert. In solving this cellular jigsaw puzzle, it is evident that sophisticated regulatory mechanisms ensure that signaling enzymes encounter their intracellular substrates in the right place and at the right time. This requires a delicate balance between two apparently opposing processes: the diffusion of information through the cytoplasm and nucleus, and the processing of information by immobilized multiprotein complexes. Cells have evolved a variety of clever ways to fulfill these requirements: (i) signal-dependent formation of protein complexes; (ii) processing of signals through preassembled multiprotein complexes; (iii) enzyme regulation by subcellular localization; and (iv) temporal control of signaling pathways. Here we highlight recent advances in our understanding of these essential regulatory processes.

Signal-Dependent Formation of Protein Complexes

Information relay from one cellular location to another often requires the dynamic formation of protein complexes. This can be initiated by posttranslational modification, switchlike functions of guanosine triphosphatases (GTPases), or protein oligomerization to generate pockets of concentrated enzyme activity. Signaling proteins typically have a modular organization, being composed of domains with binding or catalytic functions, interspersed with regions that serve as docking or substrate sites for other molecules. Currently, about 100 specialized protein-interaction modules have been identified that recognize a plethora of chemical signals. This section will compare and contrast signaling pathways that use phosphorylation, phosphoinositides, ubiquitination, and acetylation as their primary means of communication.

The Src homology 2 (SH2) domain is the archetypal protein-interaction module. Initially discovered in the $P130^{Gag\text{-}Fps}$ oncoprotein, SH2 domains consist of about 100 amino acids that bind to specific phosphotyrosine (pY)–containing peptide motifs (*1*, *2*). The human genome encodes 120 SH2 domains embedded in a variety of proteins (Fig. 1A). Most SH2 domain–containing proteins are recruited to sites of tyrosine phosphorylation to aid in the construction of molecular machines (*3*). The molecular glue in these transduction units is frequently a cohort of adaptor proteins (Fig. 1A). These noncatalytic organizing proteins contain a domain that selectively recognizes the activated receptor (an SH2 domain in the case of receptor tyrosine kinases), linked to domains such as SH3, that engage specific downstream targets, typically by binding to proline-rich sequences. Multivalent modular proteins such as Grb2 (SH3-SH2-SH3) and Nck (SH3-SH3-SH3-SH2) exemplify this strategy (Fig. 1A). For example, growth factor–induced autophosphorylation of receptor tyrosine kinases on the inner face of the plasma membrane creates pockets of phosphotyrosine. This recruits the Grb2 adaptor protein and effector proteins such as phosphatidylinositol 3-kinase (PI3K) or phospholipase Cγ (PLC-γ) to initiate downstream signaling pathways that contribute to oncogenesis and cancer cell proliferation (*4*). Accordingly, if the autophosphorylation of receptor tyrosine kinases is suppressed pharmacologically, the assembly of the downstream signaling complexes may be halted, providing a chance to slow the progression of certain cancers (*5*). Despite their apparent molecular simplicity, adaptor proteins such as Nck, which links pY signals to the actin cytoskeleton, can influence sophisticated biological processes. In neurons, Nck proteins are required for the guidance of spinal cord axons, and the formation of neuronal circuits required for walking (*6*). Nck is also required for the proper architectural organization of podocytes (specialized cells forming a filtration barrier in the kidney) and is a candidate for involvement in human diabetic nephropathy (*7*).

SH2 domains can also directly influence enzymatic activity (Fig. 1, B and C). For example, in the active configuration of the human FES cytoplasmic tyrosine kinase, the SH2 domain is integrated with the kinase domain to form a single functional unit that is only fully active when bound to a primed substrate (*8*). This nicely illustrates that the FES SH2 domain is a multifunctional device with a conventional phosphopeptide-binding site and an entirely distinct surface that stabilizes the active kinase [Fig. 1B (*8*)]. In other molecular contexts, SH2 domains suppress tyrosine kinase activity. When fused to SH3 domains, they inhibit enzymes such as Abl, Src, Lyn, and Fyn. Docking of the SH2 domain on the back face of the catalytic core allows a flexible linker to form an internal binding site for the SH3 domain (Fig. 1C). Consequently, this SH2-SH3 unit stabilizes an inactive conformation of the enzyme (*9*, *10*). Furthermore, the orientation of each SH2-SH3 unit subtly, but distinctly, shapes the topology of the substrate and adenosine 5′-triphosphate (ATP)–binding pockets in each enzyme. This latter feature may help explain the selectivity of certain ATP inhibitor drugs such as imatinib (Gleevec STY-571) that preferentially blocks Abl kinase activity to combat chronic myelogenous leukemia (*11*).

Modular domains also regulate serine or threonine phosphorylation events. This proceeds through pThr and pSer binding modules such as WW and Forkhead homology–associated (FHA) domains and 14-3-3 proteins (*12*, *13*). The transient nature of these protein-protein interactions implies that these modules must perform a balancing act. Their affinity for a particular binding site must not be too high, or binding will not be regulated by phosphorylation. Yet SH2, WW, or FHA domains must also recognize adjacent residues to allow discrimination between different phosphorylated sites. Furthermore, these modules must display sufficiently high off-rates for rapid and reversible signal transduction. Certain modules exhibit specificities for phosphoinositides phosphorylated at different sites within the inositol ring. A hallmark of these domains is the ability to target their host proteins to specific subcellular localizations, for example, through the recognition of phosphoinositides that mark particular membranes. FYVE domains frequently recognize phosphatidylinositol 3-phosphate [PI(3)P], and thereby direct proteins to endosomes, whereas pleckstrin homology (PH) domains can recognize phosphatidylinositol 4,5-bisphosphate [$PI(4,5)P_2$] or $PI(3,4,5)P_3$, which localize proteins to the plas-

[1]Department of Pharmacology, Howard Hughes Medical Institute, Box 357750, University of Washington School of Medicine, Seattle, WA 98195, USA. [2]Samuel Lunenfeld Research Institute, Mt. Sinai Hospital, 1084-600 University Avenue, Toronto, Ontario, Canada M5G 1X5. [3]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada M5S 1A8.

*To whom correspondence should be addressed. E-mail: scottjdw@u.washington.edu; pawson@lunenfeld.ca

ma membrane, often in response to PI3K signaling (*14*, *15*). A well-studied example is the protein kinase B (PKB)/Akt protein kinase, and its regulator phosphoinositide-dependent protein kinase 1 (PDK1), both of which have PH domains selective for $PI(3,4,5)P_3$ or $PtdIns(3,4)P_2$ and are thus recruited to the plasma membrane and activated in response to PI3K activity (Fig. 1A). Similarly, PLC-δ has a PH domain that selectively binds $PI(4,5)P_2$ with high affinity, which targets the enzyme to regions of membrane enriched in its phospholipid substrate (*16*, *17*). The exquisite selectivity of some PH domains for different phospholipids makes PH domain–containing proteins sensitive to the activities of enzymes that either phosphorylate or dephosphorylate these sites on the inositol ring, such as PI3K or the lipid phosphatase PTEN (phosphatase and tensin homolog deleted from chromosome 10) (*18*). Accordingly, these enzymes can modulate the localization of downstream signaling proteins that sense distinct phospholipid products. This provides an effective means of assembling or disassembling signaling complexes in different subcellular compartments.

Protein ubiquitination is also used for cell communication (*19*). This requires the enzymatic attachment of ubiquitin, a 76–amino acid protein tag, to lysine residues on the surface of target substrates (Fig. 2A). Polyubiquitin chains are formed where each ubiquitin molecule is linked through an isopeptide bond to a lysine (K) residue (such as K48, K63 or K11) within another ubiquitin molecule (Fig. 2A). Polyubiquitination often leads to the degradation of target proteins by the 26*S* proteosome. However, the ligation of mono-ubiquitin and di-ubiquitin chains also mediates other cellular functions (*20*). At least 20 structurally distinct ubiquitin-binding domains (UBDs) are embedded in a variety of proteins, which interpret information conferred by protein ubiquitination in a manner that is reminiscent of phosphotyrosine signaling (*21*). There are ~600 ubiquitin E3 ligases and more than 90 deubiquitinating proteinases encoded by the human genome. In comparison, there are 523 human protein kinases and 138 protein phosphatases (*22*). Given such comparable numbers, it is not surprising that ubiquitination and phosphorylation enzymes act synergistically in macromolecular complexes. However, there are added features in events controlled by ubiquitination. For example, UBDs generally bind ubiquitin with micromolar affinities (*23*). Consequently, these protein-protein interactions are readily reversible and bestow an element of inherent instability within the networks they assemble. This may explain why UBD proteins often control transient cellular processes such as endosomal sorting, vesicular trafficking, and events leading to autophagy (*24*, *25*). It is also worth noting that ubiquitin-like molecules such as SUMO (small ubiquitin-like modifier), Nedd8, and ISG15 are often used as covalent tags to modulate protein function and localization (*20*, *26*).

The ε-amino group of lysine can also undergo reversible acetylation, a versatile form of covalent modification used in different contexts, but often to evoke changes in the activity of histones and transcription factors [Fig. 2B; (*27*)]. These events are catalyzed by enzymes, often called histone acetyltransferases (HATs), and reversed by histone deaceylases (HDACs). In its simplest form, acetylation functions as an on/off switch to inhibit enzymes such as acetyl-CoA (coenzyme A) synthase and nitric oxide synthase [Fig. 2B; (*28*)]. Acetylation-dependent protein recruitment of bromodomains, a domain that recognizes acetylated lysine residues, is prevalent in proteins involved in chromatin remodeling (Fig. 2B). In other contexts, autoacetylation initiates dimerization of proteins with intrinsic HAT activity such as the transcriptional co-activator CBP/p300 (Fig. 2B). Finally, lysine acetylation can prevent ubiquitination of the same side-chain, which can be used to prolong the lifetime of proteins that are subject to ubiquitin-proteosome–mediated degradation (Fig. 2B). A classic example is regulation of tumor suppressor protein and transcription factor p53. Under normal conditions, p53 is polyubiquitinated and rapidly degraded by the 26*S* proteosome (*29*). However, a phosphorylation-acetylation cascade favors the stabilization of p53 in response to DNA damage. When cells are exposed to DNA damage, the phosphorylation of p53 by stress-activated kinases allows its association with p300, which in turn acetylates lysine residues to protect the tumor suppressor from ubiquitin-proteosome–mediated degradation (*30*).

Thus, distinct covalent modifications can be used in an integrated manner to facilitate the signal-dependent recruitment of proteins. Acetylation of specific lysines can also be used to prevent other major posttranslational modifications. For example, acetylation of individual lysine residues can abolish substrate recognition sites for basophilic protein kinases or, as is the case with p53, occupy ε-amino groups that would otherwise be available for ubiquitination or methylation (Fig. 2B). Thus, the initial pattern and type of posttranslational modifications determine the signaling fate of a given protein, whether it is activation, translocation, or proteosomal destruction. Undoubtedly, high-resolution mass spectrometry will prove to be the best way to explore this phenomenon. A recent study that resolved lysine acetylation profiles of 1750 proteins demonstrated that changes in the amount and frequency of this covalent modification alter a variety of signaling fates, including ubiquitin-mediated degradation and phosphorylation-dependent interactions with 14-3-3 (*31*).

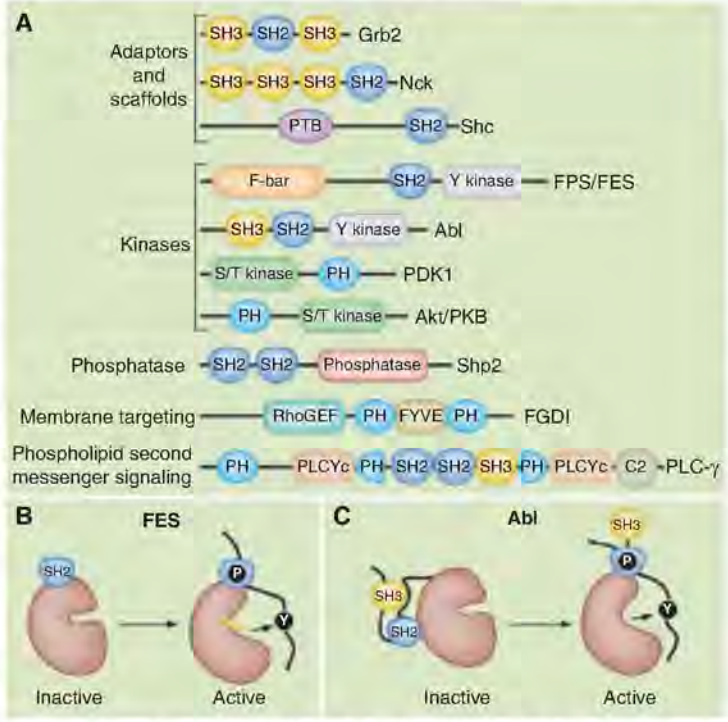


Fig. 1. The modular organization of signaling proteins. (**A**) Schematic diagram depicting the modular domain organization of selected adaptor and scaffolding proteins (Grb2, Nck, Shc); the tyrosine kinases (Fps/FES and Abl); the serine/threonine (S/T) kinases (PDK 1 and Akt/PKB); the protein tyrosine phosphatase (Shp2); a membrane-targeted Rho guanine nucleotide exchange factor (FGDI); and phospholipase C–gamma (PLC-γ). Individual protein modules are indicated: Src homology 2 (SH2); Src homology 3 (SH3); phosphotyrosine binding (PTB); FCHo2-Bin/Amphiphysin/Rvs domain (F-BAR); Fab-1, YGL023, Vps27, and EEA1 domain (FYVE); and Ca^{2+}-dependent membrane-targeting (C2) module. Enzymatic units in each protein are named. (**B**) Model depicting the active conformation of FES where the SH2 and tyrosine kinase domains form a single functional unit bound to a primed substrate (*8*). (**C**) Abl is maintained in an inactive state through the docking of the SH2 domain on the back face of the catalytic core. These intramolecular interactions are broken upon substrate binding (*11*).

Processing Signals Through Preassembled Multiprotein Complexes

The passage of signals through preassembled multiprotein signaling complexes is another means of

managing intracellular communication. Configuring enzymes in this manner not only enhances the precision of information flow but also improves the fidelity of cell signaling events by clustering successive enzymes into a transduction pathway (Fig. 3A). Commonly, intermediate enzymes in such pathways exhibit restricted substrate specificities and limited spheres of action. In fact, their only true substrate may be the next enzyme in the cascade.

This is the case for mitogen-activated protein (MAP) kinases, which form three enzyme-regulatory cascades. Extracellular signals trigger these cascades by stimulating the first member of the pathway, a MAP kinase kinase kinase (MAPKKK). Activated MAPKKKs phosphorylate mitogen-activated protein kinase kinases called MAPKKs or MEKs. This intermediary enzyme phosphorylates the mitogen-activated kinases (MAPKs). The terminal enzyme is then free to act on various downstream targets, including other protein kinases, transcriptional factors, and cytoskeletal components. In mammalian cells, scaffold proteins such as KSR and MP-1 bring together different combinations of MAP kinases to facilitate their activation and sequester these signaling units (Fig. 3A). Likewise, the linear flow of information through the Jun kinase cascade is enhanced by a family of Jun N-terminal kinase (JNK) interacting proteins called JIPs (*32*). The rationale for such an intricate means of organization lies in how signals are transferred from one enzyme to the next. For example, recruitment of the MAPKKK Raf to the KSR scaffold optimally positions the enzyme in proximity to its target substrate MEK (*33*, *34*). Phospho-MEK is then able to relay the signal to the MAPK (Erk1/2), which involves dual phosphorylation of a threonine and tyrosine in the activation loop of the MAPK. This favors the rapid dissemination of information from one enzyme to the next.

Another useful property of enzyme scaffolding is to segregate enzymes in a manner that prevents indiscriminate cross talk. This is particularly important in unicellular organisms such as yeast where mating, invasive growth, and the response to high osmolarity are regulated by distinct MAPK pathways that share a common MAPKKK called Sterile 11 (Ste11). Segregation of Ste11 activity involves binding to scaffolding proteins such as Pbs2 and Sterile 5 (Ste5). Recruitment of Ste11 into the osmosensing pathway requires interaction with Pbs2 (*35*). This chimeric protein scaffolds Ste11 and encodes its downstream target, the MAPKK. In contrast, Ste5 organizes Ste11, and the kinases Ste7 and Fus3, to direct signals through the yeast mating pathway (*36*). Ste5 also facilitates the activation of its kinase-binding partners; a regulatory domain in Ste5 that catalytically unlocks the Fus3 kinase for phosphorylation by Ste7 (*37*). Thus, Ste5 not only organizes successive components of a yeast MAPK cascade but allosterically modifies the conformation of its bound kinases, making them more amenable to activation (*37*).

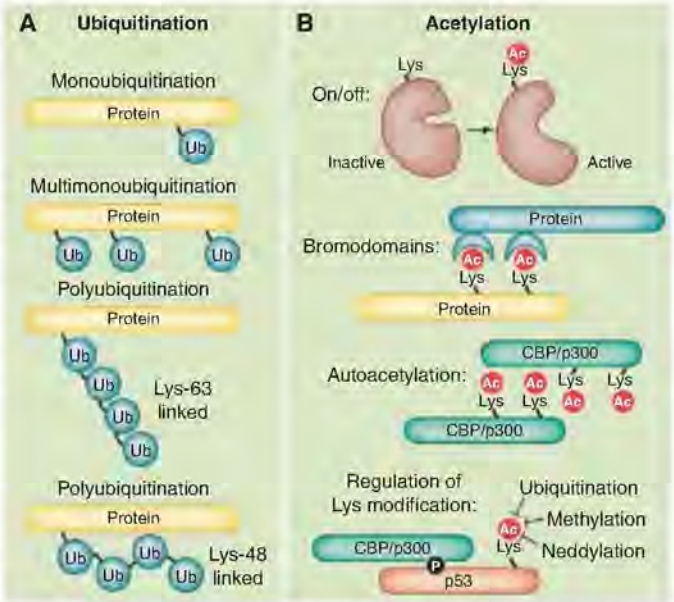


Fig. 2. Regulation by ubiquitination or acetylation. These covalent modifications are frequently used for cell communication. **(A)** Ligation of ubiquitin (Ub) can occur in different patterns on proteins. **(B)** Acetylation in different contexts serves as an enzymatic on/off switch; for recruitment of bromodomain proteins; (autoacetylation) to induce protein-protein interactions; or for regulation of further lysine modifications.

Enzyme Regulation by Subcellular Localization

Compartmentalization of enzymes in proximity to substrates is another means of spatially restricting cell-signaling events. Accordingly, a plethora of kinase- and phosphatase-binding proteins tether their enzyme-binding partners to sites where they can preferentially receive activating signals and be close to selected substrates. Prime examples remain the type 1 phosphatase (PP1) targeting subunits and A-kinase anchoring proteins (AKAPs). The concept of phosphatase targeting was proposed about the same time as the kinase-anchoring hypothesis as a means to generate substrate specificity for second messenger–regulated phosphorylation events (*38*). The glycogen-particle associated protein GM was the first PP1-targeting subunit to be identified. GM and a functionally related molecule called PTG coordinate signaling complexes that influence glycogen metabolism. Modulation of targeted PP1 activity involves a K/R-V/I-X-F (F, Phe; I, Ile; K, Lys; R, Arg; V, Val; and X, any amino acid) motif in GM that inserts into a groove distal to the active site of the enzyme (*39*). Peptides encompassing this region are sufficient to displace the PP1 catalytic subunit from GM and abolish the preferential dephosphorylation of glycogen-associated substrates. Thus, GM not only targets PP1 but also allosterically regulates phosphatase activity. The RVxF motif has now been identified in more than 50 potential PP1-targeting subunits. As an example, the muscle-specific phosphatase holoenzyme (PP1-M) contains a targeting subunit called M110/MBS that directs phosphatase activity toward a select group of muscle proteins, including myosin and possibly moesin. Additionally, M110/MBS nucleates a signaling complex with the guanosine 3′,5′-monophosphate (cGMP)–dependent kinase (PKG) and the Rho GTPase. Mobilization of the second messenger cGMP activates PKG that, in turn, phosphorylates M110/MBS to trigger events that lead to smooth-muscle relaxation (*40*).

Spatial organization of the cAMP-dependent protein kinase (PKA) holoenzyme [consisting of a regulatory (R) subunit dimer and two catalytic (C) subunits] is achieved through interaction with AKAPs (Fig. 3B). High-affinity interaction with PKA is mediated by an amphipathic helix on the AKAP that inserts into a hydrophobic pocket formed by the R-subunit dimer (*41*, *42*). Targeting determinants within the anchoring protein confer the subcellular localization of PKA-AKAP complexes to specific organelles. Mammalian genomes encode about 20 AKAP genes that generate ~75 alternately spliced transcripts. Consequently, multiple variants and differentially targeted isoforms of the same anchoring protein are often expressed within the same cell (Fig. 3C). This increases the repertoire of intracellular PKA anchoring sites and provides a means to restrict action of this broad-spectrum protein kinase toward only a few of its potential substrates. Some AKAPs can also interface with other cAMP signaling elements including adenylyl cyclases, phosphodiesterases, and Epac guanine nucleotide exchange factors (*43*, *44*). Live-cell imaging approaches have shown that these anchored phosphodiesterases participate in negative-feedback loops that locally suppress adenosine 3′,5′-monophosphate (cAMP). Accordingly, these cAMP-responsive units generate local fluctuations in cAMP and concomitant pulses of PKA or Epac activity. Thus, AKAPs appear to orchestrate all aspects of cAMP signaling (Fig. 3B).

A broader role for AKAPs in the spatial organization of cell signaling events became apparent when it was shown that they interact with other regulatory enzymes (Fig. 3B). For example, AKAP79/150 interacts with PKA, the calcium/

phospholipid-dependent protein kinase C (PKC) and calmodulin-dependent phosphatase PP2B (*45*). This implies that second messenger signals that control the phosphorylation and second messenger signals that favor the dephosphorylation of a target substrate pass through the same AKAP complex. This type of regulation may be particularly important for the control of rapid signaling events such as the modulation of neuronal ion channels (Fig. 3C). Strategic use of distinct anchored enzyme combinations provides another way to expand the repertoire of cellular events that the same AKAP modulates. For example, different enzyme combinations anchored to AKAP79/150 modulate the activity of two neuronal ion channels: AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid)–type glutamate receptors and muscarine-sensitive potassium channels (*46*). In hippocampal neurons, AKAP79/150 coordinates PKA and PP2B modulation of AMPA currents, while any AKAP79/150-associated PKC remains inactive in this process. In contrast, AKAP79/150 enables PKC to facilitate M-current regulation in SCG (superior cervical ganglia) neurons, while PKA and PP2B appear to be nonessential (*46*).

Unlike scaffolding proteins, which process information in a linear manner (Fig. 3A), the combinatorial assembly of AKAP-enzyme complexes permits the integration and dissemination of multiple signals (Fig. 3B). Although the contextual cues that drive the preferential assembly of distinct AKAP complexes are unclear, one possibility is that the initial binding event of the anchoring protein with its target substrate promotes a succession of conformational changes that select the recruitment of the next binding partners. However, cotranslational assembly of protein complexes through localized protein synthesis, species-specific or cell type–specific expression of particular binding partners may further influence the composition of these "context-dependent" signaling networks.

Temporal Control of Signaling Pathways

Changes in composition or the amount of enzyme complexes over time also modulate cellular events. This often involves phosphorylation, ubiquitin-mediated degradation, and translocation of signaling components. Nuclear factor κB (NF-κB) is a transcription factor that regulates expression of genes involved in inflammation, apoptosis, and tumorigenesis. Activation of the NF-κB involves the phosphorylation and ubiquitination of several components [Fig. 4A; (*47*)]. In the absence of stimuli, NF-κB is kept from the nucleus by its interaction with the inhibitory partner IκB. Upon stimulation, IκB is phosphorylated, resulting in subsequent ubiquitination and degradation. The degradation exposes a nuclear localization sequence (NLS) on NF-κB, favoring its translocation to the nucleus. The IκB kinase complex, composed of two catalytic subunits (Iκκα and Iκκβ) and the regulatory subunit (Iκκγ/NEMO), regulates the phosphorylation of IκB (*48*). Stimulation of the NF-κB pathway results in the ubiquitination of NEMO and activation of the catalytic subunits (Fig. 4A). The enzyme then phosphorylates IκB, resulting in the recruitment of the E3 ubiquitin ligase SCF-βTRCP. This promotes Lys^{48}-linked ubiquitination of IκB, leading to its degradation (*49*). Because ubiquitination plays a central role in NF-κB activation, its removal by deubiquitinating enzymes is critical to the down-regulation of the NF-κB signal. The cylindromatosis tumor suppressor protein (CYLD) removes Lys^{63}-linked chains from several NF-κB pathway members, including Iκκγ/NEMO, to regulate the duration of NF-κB activation (*22*). Likewise, dephosphorylation of the IκB kinase complex by protein phosphatase 2A further attenuates the NF-κB response (*47*). Thus, synchronized ubiquitination and phosphorylation events can exert precise temporal control on a signaling pathway. Yet, the expression of NF-κB target genes often occurs 4 to 6 hours after agonist stimulation. This suggests that a rate-limiting step appears to be the time it takes for NF-κB to translocate from the cytoplasm into the nucleus.

In oxygen-sensitive tissues, transcriptional responses can occur in minutes rather than hours. The hypoxia-inducible factor 1α (HIF-1α) is rapidly induced in cardiomyocytes and certain tumors in response to reduced intracellular oxygen (*50*). Under normoxic conditions, the abundance of HIF-1α is kept low through its ubiquitin-mediated proteosomal degradation. This process is initiated by the hydroxylation of two conserved proline residues (Pro^{402} and Pro^{564}) in HIF-1α, by oxygen-sensitive prolyl hydroxylases (PHDs) [Fig. 4B; (*51*)]. The hydroxylated proline residues constitute a binding site for the von Hippel–Lindau protein (pVHL), which is part of a complex that ubiquitinates HIF-1α and targets it for degradation by the proteosome (*52*). During hypoxia, the continual destruction of HIF-1α is halted by the enzymatic activity of PHD that ceases in the absence of oxygen, and a ubiquitin E3 ligase

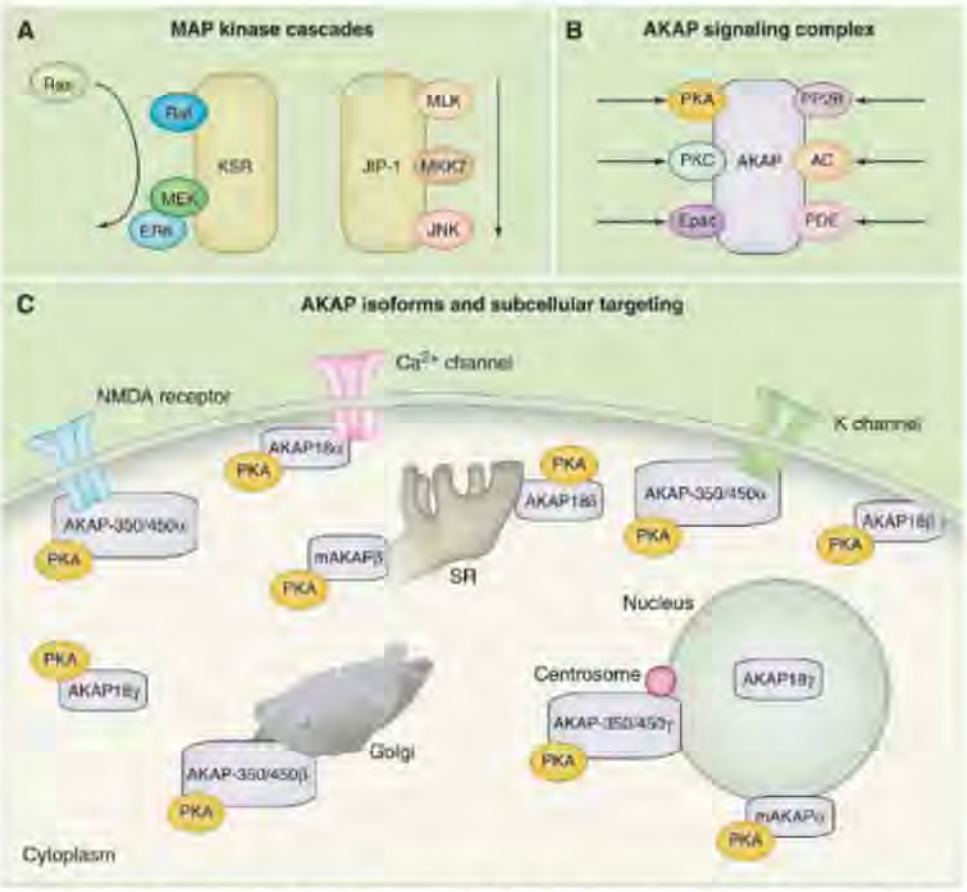


Fig. 3. Signal transduction through preassembled enzyme complexes. (**A**) The linear flow of information through the MAP kinase cascades organized by kinase suppressor of Ras (KSR) and JNK interacting protein (JIP). (**B**) A-kinase anchoring proteins (AKAPs) organize protein kinase A (PKA), guanine nucleotide exchange factors (Epac) and phosphodiesterases (PDE) into cAMP-responsive complexes. Anchoring of certain protein phosphatases (PP2B) and PKC broaden AKAP function. (**C**) PKA anchoring to various AKAP isoforms targets the kinase to defined subcellular locations. Diagram of a prototypic cell showing the targeting of PKA via AKAP18, AKAP350/450, and mAKAP variants. NMDA, *N*-methyl-D-aspartate.

(seven in absentia homolog 2, Siah2) that ubiquitinates PHDs and targets them for proteosomal degradation [Fig. 4B; (*53*)]. Collectively, these mechanisms terminate the continual destruction of HIF-1α, which allows the protein to form a stable heterodimeric complex with the HIF-1β subunit. The HIF-1α–HIF-1β heterodimer accumulates in the nucleus and initiates transcription of proangiogenic, metabolic, and antiapoptotic genes that promote cell survival. Recent evidence suggests that the anchoring protein (mAKAP) organizes ubiquitin E3 ligases that manage the stability of HIF-1α (*54*). In cardiomyocytes, depletion of mAKAP or disruption of its targeting to nesprin, a protein that forms the outer ring of the nuclear pore complex, alters the stability of HIF-1α and activation of genes associated with hypoxia. Anchoring of an oxygen-sensitive, ubiquitin-mediated destruction complex at the nuclear pore optimizes temporal control HIF-1α to suppress the hypoxic response. Yet when a hypoxic environment prevails, the transcription factor is released from this anchored complex and immediately translocates to its site of action inside the nucleus (Fig. 4B).

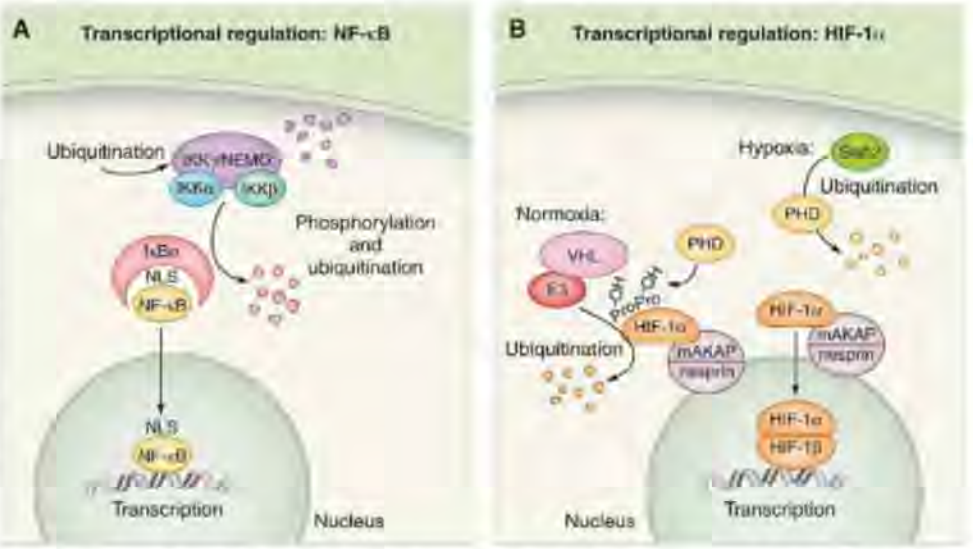


Fig. 4. Temporal regulation of signaling events. (**A**) Schematic diagram of the NF-κB pathway; individual protein components are indicated. (**B**) The concerted actions of oxygen-sensing enzymes and ubiquitin E3 ligases control the stability of the hypoxia-inducible factor protein HIF-1α. Schematic diagram of the HIF-1α pathway under normoxic (left) and hypoxic (right) conditions; individual protein components are indicated.

Where Is This Taking Us?

In the past decade, we have witnessed unparalleled advances in our understanding of cell signaling. These include the advent of kinase inhibitor drugs; an appreciation of how phosphorylation, ubiquitination, and acetylation events instigate protein-protein interactions; and the realization that enzyme compartmentalization determines signaling specificity. Technological advances in mass spectrometry, high-throughput genomic sequencing, genetically encoded fluorescent proteins, RNA interference, and live-cell imaging have helped us understand how cellular information is resolved in space and time. If a desirable goal is to deftly manipulate the enzyme activity in space and time, then the future may lie in the burgeoning field of synthetic biology. By taking advantage of the existing knowledge of modular domains, investigators are already generating synthetic molecules that redirect signaling in situ. Prime examples include artificial guanine nucleotide exchange factors designed to "rewire" actin reorganization and alter cell morphology (*55*), or the design of light-activated Rac1 GTPases that permit laser-induced membrane ruffling at any point in the cell (*56*). Undoubtedly, these and other innovative approaches will assist our ongoing molecular exploration of the cell.

References and Notes

1. I. Sadowski, J. C. Stone, T. Pawson, *Mol. Cell. Biol.* **6**, 4396 (1986).
2. B. J. Mayer, M. Hamaguchi, H. Hanafusa, *Cold Spring Harb. Symp. Quant. Biol.* **53**, 907 (1988).
3. Z. Songyang *et al.*, *Mol. Cell. Biol.* **14**, 2777 (1994).
4. H. Kouhara *et al.*, *Cell* **89**, 693 (1997).
5. A. Ostman, C. H. Heldin, *Adv. Cancer Res.* **97**, 247 (2007).
6. J. P. Fawcett *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 20973 (2007).
7. N. Jones *et al.*, *Nature* **440**, 818 (2006).
8. P. Filippakopoulos *et al.*, *Cell* **134**, 793 (2008).
9. I. Moarefi *et al.*, *Nature* **385**, 650 (1997).
10. M. A. Young, S. Gonfloni, G. Superti-Furga, B. Roux, J. Kuriyan, *Cell* **105**, 115 (2001).
11. B. J. Druker *et al.*, *N. Engl. J. Med.* **344**, 1038 (2001).
12. M. B. Yaffe *et al.*, *Cell* **91**, 961 (1997).
13. M. B. Yaffe, L. C. Cantley, *Nature* **402**, 30 (1999).
14. A. Hayakawa *et al.*, *J. Biol. Chem.* **279**, 5958 (2004).
15. M. A. Lemmon, *Nat. Rev. Mol. Cell Biol.* **9**, 99 (2008).
16. D. R. Alessi *et al.*, *Curr. Biol.* **7**, 261 (1997).
17. D. A. Cross, D. R. Alessi, P. Cohen, M. Andjelkovich, B. A. Hemmings, *Nature* **378**, 785 (1995).
18. J. O. Lee *et al.*, *Cell* **99**, 323 (1999).
19. A. Ciechanover, S. Elias, H. Heller, S. Ferber, A. Hershko, *J. Biol. Chem.* **255**, 7525 (1980).
20. C. Grabbe, I. Dikic, *Science* **322**, 872 (2008).
21. L. Hicke, H. L. Schubert, C. P. Hill, *Nat. Rev. Mol. Cell Biol.* **6**, 610 (2005).
22. M. E. Sowa, E. J. Bennett, S. P. Gygi, J. W. Harper, *Cell* **138**, 389 (2009).
23. S. Polo *et al.*, *Nature* **416**, 451 (2002).
24. V. Kirkin, D. G. McEwan, I. Novak, I. Dikic, *Mol. Cell* **34**, 259 (2009).
25. C. Raiborg, H. Stenmark, *Nature* **458**, 445 (2009).
26. J. H. Hurley, S. Lee, G. Prag, *Biochem. J.* **399**, 361 (2006).
27. A. Basu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13785 (2009).
28. B. D. Strahl, C. D. Allis, *Nature* **403**, 41 (2000).
29. B. Vogelstein, D. Lane, A. J. Levine, *Nature* **408**, 307 (2000).
30. N. A. Barlev *et al.*, *Mol. Cell* **8**, 1243 (2001).
31. C. Choudhary *et al.*, *Science* **325**, 834 (2009).
32. A. J. Whitmarsh, J. Cavanagh, C. Tournier, J. Yasuda, R. J. Davis, *Science* **281**, 1671 (1998).
33. J. Muller, S. Ory, T. Copeland, H. Piwnica-Worms, D. K. Morrison, *Mol. Cell* **8**, 983 (2001).
34. T. Rajakulendran, M. Sahmi, M. Lefrancois, F. Sicheri, M. Therrien, *Nature* **461**, 542 (2009).
35. F. Posas, H. Saito, *Science* **276**, 1702 (1997).
36. K.-Y. Choi, B. Satterberg, D. M. Lyons, E. A. Elion, *Cell* **78**, 499 (1994).
37. M. Good, G. Tang, J. Singleton, A. Remenyi, W. A. Lim, *Cell* **136**, 1085 (2009).
38. P. Cohen, T. W. Cohen, *J. Biol. Chem.* **264**, 21435 (1989).
39. M. P. Egloff *et al.*, *EMBO J.* **16**, 1876 (1997).
40. H. K. Surks *et al.*, *Science* **286**, 1583 (1999).
41. M. G. Gold *et al.*, *Mol. Cell* **24**, 383 (2006).
42. F. S. Kinderman *et al.*, *Mol. Cell* **24**, 397 (2006).
43. K. L. Dodge-Kafka *et al.*, *Nature* **437**, 574 (2005).
44. A. L. Bauman *et al.*, *Mol. Cell* **23**, 925 (2006).
45. T. M. Klauck *et al.*, *Science* **271**, 1589 (1996).
46. N. Hoshi, L. K. Langeberg, J. D. Scott, *Nat. Cell Biol.* **7**, 1066 (2005).
47. M. Karin, Y. Cao, F. R. Greten, Z. W. Li, *Nat. Rev. Cancer* **2**, 301 (2002).
48. F. Mercurio *et al.*, *Science* **278**, 860 (1997).
49. S. Rahighi *et al.*, *Cell* **136**, 1098 (2009).
50. B. Keith, M. C. Simon, *Cell* **129**, 465 (2007).
51. E. Berra *et al.*, *EMBO J.* **22**, 4082 (2003).
52. D. J. George, W. G. Kaelin Jr., *N. Engl. J. Med.* **349**, 419 (2003).
53. K. Nakayama *et al.*, *Cell* **117**, 941 (2004).
54. W. Wong, A. S. Goehring, M. S. Kapiloff, L. K. Langeberg, J. D. Scott, *Sci. Signal.* **1**, ra18 (2008).
55. C. J. Bashor, N. C. Helman, S. Yan, W. A. Lim, *Science* **319**, 1539 (2008).
56. Y. I. Wu *et al.*, *Nature* **461**, 104 (2009).

57. We acknowledge C. Pawson and M. Milnes for evaluation of this manuscript and L. Langeberg for design of the figures. J.D.S. is supported by NIH grant GM48231 and Leducq transatlantic network 06CVD02. T.P. is supported by the National Cancer Institute of Canada, Canadian Cancer Society, Canadian Institutes for Health Research (grants 6849 and 57793), Genome Canada, and a Terry Fox Program Project grant.

10.1126/science.1175668

REVIEW

Why and How Bacteria Localize Proteins

L. Shapiro,[1]* H. H. McAdams,[1] R. Losick[2]

Despite their small size, bacteria have a remarkably intricate internal organization. Bacteria deploy proteins and protein complexes to particular locations and do so in a dynamic manner in lockstep with the organized deployment of their chromosome. The dynamic subcellular localization of protein complexes is an integral feature of regulatory processes of bacterial cells.

Bacteria were once viewed as amorphous reaction vessels with chromosomes that wandered freely and randomly throughout the cell. The advent of genetically encoded fluorescent reporters harnessed to powerful cell-imaging technologies has enabled in vivo tracking of protein movement and revealed a strikingly complex inner world within bacteria. This inner environment is exquisitely organized, in a highly controlled state of flux, and responsive to changing functions demanded of the cell. For example, some proteins oscillate rapidly from one end of the cell to the other, whereas others form dynamic helices along the length of the cell or rings across its midsection, and yet others form distributed focal complexes on the cell's surface or clusters at specific intracellular sites. Processes controlled at multiple levels construct (and remove) subsystems and surface structures at specific times and places in response to internal and external signals. This dynamic internal architecture facilitates behaviors as diverse as symmetric and asymmetric division, motility, chemotaxis, morphological differentiation, assembly into multicellular communities, and interactions with animal and plant hosts. In one bacterial species, *Caulobacter crescentus*, at least 10% of predicted encoded proteins exhibited specific subcellular organization (*1*). Here, we explore why bacteria dynamically deploy key regulatory proteins to particular sites in the cell and how this positioning and repositioning is achieved.

Why Are Proteins Localized?

Polar positioning of chemotaxis arrays. Since the first observation that bacterial chemoreceptors, along with CheA histidine kinases and CheW adapter proteins, are located at or near one cell pole of *C. crescentus* (*Caulobacter* hereafter) and *Escherichia coli* (*2*, *3*), fluorescence microscopy and cryoelectron microscopy images have revealed the exquisite architecture of these polar complexes (*4*, *5*). The chemotaxis sensor system controls the activity of the flagellar motor so that cells move toward attractants and away from repellents. The chemoreceptor array comprises thousands of receptors arranged in a "trimer of dimers" configuration (*6*–*8*). The unit cell of this hexagonal lattice is formed by three receptor dimers (*5*). Why has the cell evolved this elaborate localized array? One proposal is that the close spacing of the components of the chemosensory array promotes signal amplification (*9*). The *Caulobacter* chemoreceptor array is always positioned somewhat away from the pole on the convex side of the crescent-shaped cell (Fig. 1B) (*5*), whereas a linear array of crescentin intermediate filaments (which confer on *Caulobacter* its distinctive shape) is always positioned on the concave side of the crescent (Fig. 1B) (*10*). Thus, the *Caulobacter* cell has dorsal-ventral asymmetry as well as anterior-posterior asymmetry exhibited by the differential polar placement of the flagellum and stalk, the chromosomal origin complex, and signaling kinases (Fig. 1A).

Chromosome organization. Subcellular protein localization also mediates the highly organized deployment of bacterial chromosomes. Bacterial DNA, if stretched out linearly, would be about 1000 times the length of the cell. To deal with these spatial constraints, bacteria have evolved a highly ordered deployment of the chromosome within the cell. Notably, the chromosome has a specific orientation within the cell. For many bacteria, including *Caulobacter*, *Vibrio cholerae* chromosome I, and sporulating *Bacillus subtilis*, the DNA sequence around the replication origin is positioned at the cell pole (*11*–*13*). In *Caulobacter*, upon replication of the chromosomal origin region, the origin-proximal *parS* centromeric sequence bound to the ParB partition protein is transported rapidly across the long axis of the cell and is captured by a polar polymeric network of the PopZ protein (Fig. 2A) (*14*, *15*). In growing cells of *B. subtilis*, origins are located at the outer edge of the nucleoids. But in cells that have entered the pathway to sporulation, a hemispheric array of the DivIVA protein at the cell pole captures the RacA protein decorating the origin-proximal portion of the newly duplicated chromosome (Fig. 2B) (*16*, *17*). In addition to specific cellular positioning of the origin and terminus sequences, the entire chromosome is organized within the cell. When positions of fluorescently tagged foci in the cell were measured in more than 100 *Caulobacter* strains, each with a different tagged locus (*18*), a linear correspondence was observed between the position of a given locus along the length of the cell and its position on the chromosome. Analysis of the *B. subtilis* and *E. coli* genome organization also suggested a linear correspondence of gene po-

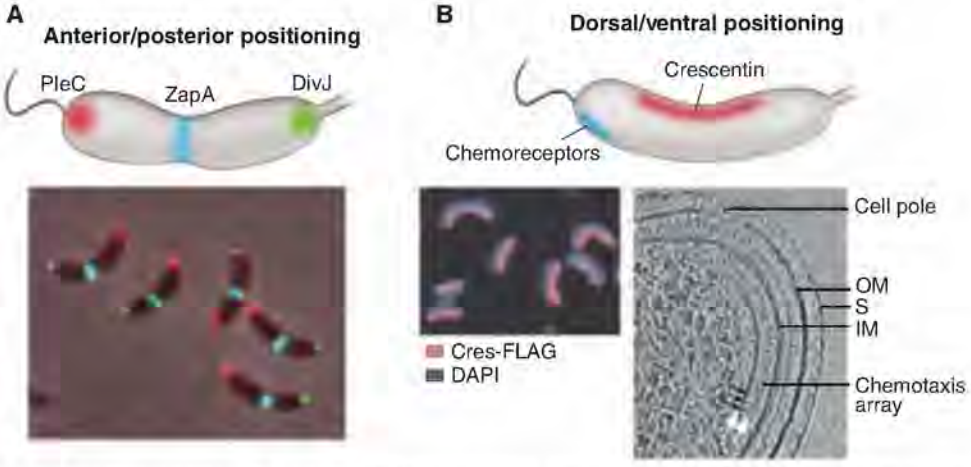


Fig. 1. Localization patterns. **(A)** An anterior-posterior cellular localization axis is exhibited by the *Caulobacter* histidine kinases PleC (red) and DivJ (green) that dynamically and selectively localize to specific cell poles. The ZapA cell division protein (blue) localizes to the FtsZ ring. **(B)** In these same cells, a dorsal-ventral localization axis is exhibited by crescentin (cres) intermediate filaments that localize along the inner concave side of the crescent-shaped cell and are responsible for this distinctive cell shape (*10*); in contrast, the chemotaxis sensor array localizes at the convex outer side of the crescent near the cell pole (*5*). DAPI, 4′,6-diamidino-2-phenylindole.

[1]Department of Developmental Biology, Stanford University, Stanford, CA 94305, USA. [2]BioLabs, Harvard University, Cambridge, MA 02138, USA.

*To whom correspondence should be addressed. E-mail: shapiro@stanford.edu

sition in the cell and gene order on the chromosome (*19*–*21*).

Cell division. When cells divide, a key challenge is to ensure that cytokinesis commences at the right time and at the right place along the cell length. How is this achieved? In many rod-shaped bacteria, assembly of the cell division ring is guided to the center of the cell by the action of inhibitors of the polymerization of the tubulin-like protein FtsZ that drives cytokinesis (Fig. 2). These inhibitors are dynamically localized to the cell poles to create an inhibitor-free zone at the cell center. In the case of *E. coli*, this process involves rapid pole-to-pole oscillation of the MinCD inhibitor complex, so that the cell center is the site of lowest average inhibitor concentration, thus allowing midcell assembly of FtsZ (Fig. 3A) (*22*). The MinE protein restricts the MinCD membrane-associated complex to the cell poles. In *B. subtilis*, which lacks MinE, MinCD is recruited to sites of septation and diffuses away from completed septa, preventing septation from occurring near new cell poles (*23*). The MinJ protein tethers MinCD to DivIVA, which is positioned at septa and in a hemispheric array at the poles (*24*, *25*). In both cases, nucleoid occlusion, in which the FtsZ ring is precluded from assembling over the chromosome, augments the action of the polar inhibitors (*26*–*29*). *Caulobacter*, in contrast, takes advantage of the concurrent initiation of DNA replication and polar segregation of the chromosomal origin to establish a gradient of the MipZ cytokinesis-inhibiting protein emanating from the cell poles, thus determining the midcell placement of the FtsZ ring (Fig. 2B) (*30*). MipZ, which interacts with ParB bound to the *parS* centromere anchored at the cell poles, inhibits the polymerization of FtsZ. The dynamic interaction of MipZ and ParB results in a gradient of MipZ, with the highest concentration at the cell poles. As the newly duplicated centromere is moved rapidly along the long axis of the cell to the opposite pole, FtsZ localizes to the site of lowest concentration of the MipZ inhibitor. The division plane is established at this point. The use of protein gradients emanating from cell poles is not restricted to bacterial cells. The fission yeast *Schizosaccharomyces pombe* establishes a gradient of polar Pom1 kinases to tie the onset of mitosis to cell length (Fig. 3C) (*31*, *32*).

These examples illustrate how different cell types use the cell's instantaneous topology and internal organization to coordinate spatially and temporally controlled events crucial to cell cycle progression.

Using geometry to control proteolysis and drive membrane movement. Protease complexes that clear proteins from the cell at the right time and place are key members of the cell's regulatory tool kit. With impressive destructive power, these proteases are held in tight control. Somewhat surprisingly, in bacterial cells one of the control mechanisms uses the dynamic positioning of protease complexes to the cell poles (*33*, *34*). This spatial control of proteolysis is seen most clearly in *Caulobacter*, where it is intimately linked to cell cycle progression (*33*).

Entry into S phase is a critical event in all cells, and no less so in bacteria. Chromosome replication in *Caulobacter* is restricted to once per cell cycle, and the decision to enter S phase is governed in part by the proteolysis of CtrA (*33*, *35*), an inhibitor of replication initiation (*36*). Clearance of CtrA from the cell allows DnaA-mediated replication initiation. CtrA, a substrate of the essential protease ClpXP (*37*), is degraded only in the stalked cell. Because the ClpXP protease is present throughout the cell cycle and has many other substrates, the cell is faced with the challenge of providing both temporal and spatial specificity to ClpXP proteolysis of CtrA. Specificity is provided, in part, by localizing both the protease complex and its CtrA substrate to the stalked cell pole at specific times in the cell cycle (*33*). The time in the cell cycle when these protein complexes are localized to the pole is governed by a dynamically localized phospho-signaling pathway and a second messenger system (*38*, *39*). The culmination of these events is assembly of a polar complex in which the ClpXP proteolytic machinery and its substrate are brought together at just the right time to clear CtrA from the cell so that DNA replication can be initiated. This use of polar positioning to control proteolysis is also observed in the soil bacterium *Sinorhizobium meliloti*, where protease polar localization is essential for coordinating cell cycle progression and for differentiation into bacteroids upon infection of plant roots (*40*).

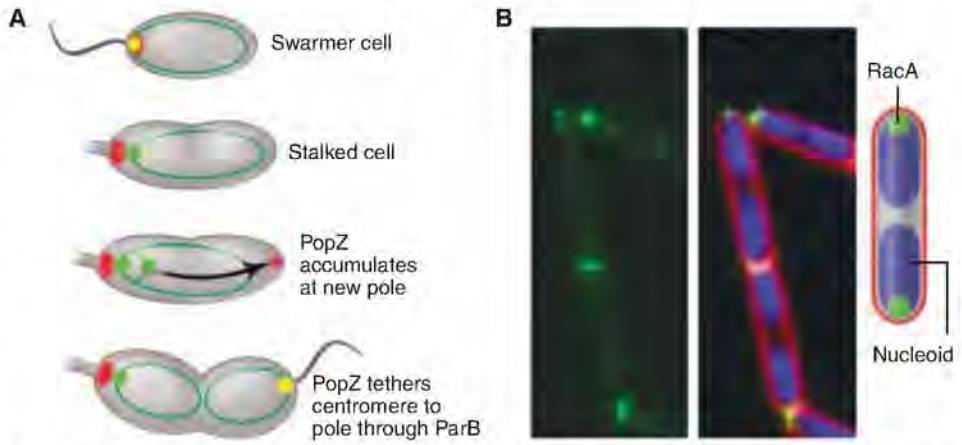


Fig. 2. Chromosome attachment to the cell pole. (**A**) The *Caulobacter parS* centromere bound to the ParB partition protein is attached to the cell pole by interaction with the polar PopZ polymeric network (*14*, *15*). The initiation of replication triggers the assembly of PopZ at the opposite pole, where it captures the duplicated copy of *parS*/ParB. The diagram shows PopZ (red), ParB (green), PopZ + ParB (yellow), and chromosomes depicted as rings (dark green). (**B**) Sporulating cells of *B. subtilis* anchor chromosomes to the cell poles via the sporulation protein RacA, which binds to sites near the replication origin and to DivIVA at the cell poles (*16*, *17*). The images show RacA tagged with green fluorescent protein (green), the nucleoids (blue), and the cell membrane (red).

How Are Proteins Localized?

Diffusion and capture. Proteins are localized to their correct position within the bacterial cell by signals encoded in their primary amino acid sequence. But how do these signals work? Eukaryotic cells have sorting machinery based on vesicles that deliver protein cargos to their proper destination, but bacteria lack vesicle-mediated sorting. Instead, bacterial protein localization is mediated principally by diffusion and capture. In the small bacterial cell, proteins can diffuse rapidly throughout the cytoplasm in three dimensions or throughout the membrane in two dimensions, encountering other proteins in all possible configurations. When the protein encounters another protein to which it specifically adheres, that configuration will tend to persist. Thus, a localized "target" protein complex can capture and localize individual proteins or groups of proteins. Depending on the energetics of the protein-protein binding, the localization interaction may be persistent or transient; in the latter case, there can be continuous exchange between the cytoplasmic and bound-protein state. Overexpression of membrane-bound docking proteins can result in their presence throughout the membrane even though the wild-type cell exhibits specific foci, suggesting saturation of the target complex docking sites in the overexpression strain (*41*).

Direct evidence in support of the diffusion-and-capture mechanism has been obtained for the localization of the PleC histidine kinase to

the cell pole in *Caulobacter* and the assembly of the BofA-SpoIVFA-SpoIVFB protease complex in sporulating *B. subtilis*. Observation of the movement of single fluorescently tagged PleC molecules in live cells revealed a stationary PleC focus at the cell pole while PleC molecules away from the pole defused in the membrane without a directional bias, supporting the hypothesis that PleC diffuses randomly until captured by a target at the pole (*42*). In sporulating *B. subtilis*, SpoIVFA is the anchor that captures BofA and SpoIVFB. A diffusion-and-capture mechanism was deduced from the observation that when membrane containing the target protein becomes topologically isolated from the cytoplasmic membrane after completion of the phagocytic-like engulfment process, BofA and SpoIVFB in the cytoplasmic membrane can no longer reach SpoIVFA (*43*). This finding demonstrated that localization required an uninterrupted membrane through which BofA and SpoIVFB can diffuse to the SpoIVFA target and be captured.

Ultimate cues. As we have seen, proteins often localize by binding to another protein or proteins that are already sequestered at a particular location in the cell. This dependency raises the question of the prior cue that dictates the position of the target protein in the ensemble of proteins, and the one before that and so on, until the position of the first localized protein is established. Consider again the cytokinetic FtsZ ring, which in *Caulobacter* assembles only at a site near the midpoint between the cell poles because the MipZ inhibitor of FtsZ polymerization interacts with the chromosomal origin bound to the cell poles to form a gradient of the inhibitor. In this case, therefore, confinement of the FtsZ ring to the cell middle ultimately derives from the organization of the chromosome and the placement of origins at the poles.

Dynamic self-assembly appears to contribute to the clustering of proteins at subcellular locations. As discussed earlier, chemoreceptors form large polar clusters. In *E. coli*, in which large clusters are seen at both poles, receptors stochastically self-assemble into small clusters all around the cell, including along the side walls, in an exponential distribution with the largest clusters at the poles (*44*). According to this self-assembly model, receptor molecules are most likely to be captured by existing clusters if one is nearby, and to nucleate new clusters otherwise. As a consequence, and in steady state, larger clusters will tend to assemble as far apart from each other as possible—that is, at the poles.

What other ultimate cues dictate protein localization in bacteria? A different type of cue has emerged from studies of two peripheral membrane proteins in *B. subtilis*: SpoVM and DivIVA. Spore formation takes place in a sporangium that consists of a cell-within-a-cell generated by an engulfment process, as noted above. The outer cell—the mother cell—nurtures the developing spore contained within it (Fig. 4). The SpoVM protein is produced in the mother cell but localizes specifically to the outer membrane surface of the developing spore rather than the cytoplasmic membrane that surrounds the mother cell. How does SpoVM discriminate between these two membranes? Multiple lines of evidence indicate that SpoVM preferentially adheres to membrane with positive (convex) curvature: Only the surface of the spherical spore is positively curved; the cytoplasmic membrane is concave (*45*). So the position of this protein—and, in turn, the position of other proteins that it recruits to the outer surface of the spore—is determined by geometry, specifically the shape of the spore membrane.

Conversely, negative membrane curvature seems to be the ultimate cue for DivIVA, a multifunctional protein that contributes to Z-ring positioning during growth (by recruiting MinCD at the *B. subtilis* cell poles) and localizing chromosome origin regions at the poles during sporulation (by capturing chromosome-bound RacA). DivIVA preferentially localizes as a disk around the outer edge of the developing septum and at the poles. Yet apparently it does not interact with any divisome or polar protein. What does it recognize? The poles and the junction of the septum to the lateral wall of the cell are sites of more extreme negative curvature than the sides of the cell, which are also negatively curved. Evidence suggests that DivIVA localizes in a hierarchical fashion favoring the most extreme concavity of the septum, next the hemispherical curvature of the inner surface of the poles, and finally the more gently curved inner surface of the sides of the cell (*46*, *47*). These findings raise the possibility that geometry may play a widespread role in protein localization in bacteria. Perhaps other proteins also localize principally to the cell ends by preferring hemispherical curvature.

Some protein localization may depend on a never-ending cyclical cascade of cues from one cell generation to the next. Cell division gives rise to the new pole of the resulting daughter cells, so certain proteins that localize to the septum

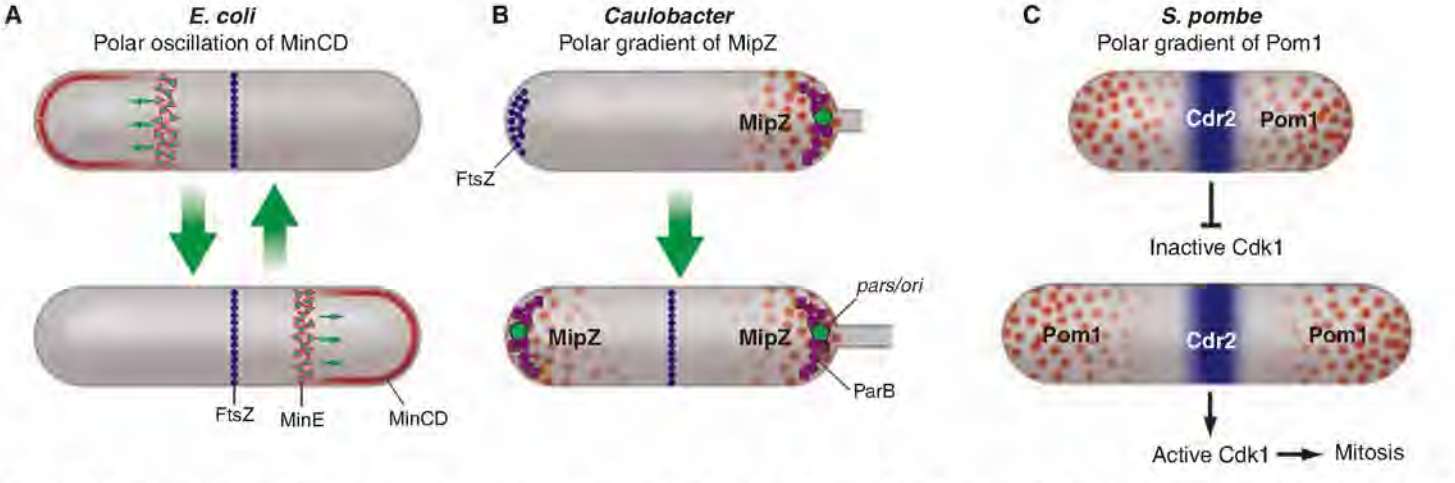


Fig. 3. Cell topology and inhibitor gradients control place and time of cell division. Repression of FtsZ polymerization by polar localized proteins that exhibit a minimum of the repressor at midcell restricts the site of division ring assembly in both *E. coli* and *Caulobacter*. However, *E. coli*'s strategy depends on oscillation of the MinCD repressor from pole to pole (**A**), whereas *Caulobacter* establishes a gradient of the MipZ repressor with the highest concentration at the cell poles (**B**). In *S. pombe*, the Pom1 repressor is localized to the poles (**C**). In small cells, the gradient of Pom1 extending from the poles overlaps at Cdr2 located at midcell and represses its activity. Consequently, the Cdk1 pathway is blocked, preventing entry into mitosis. As the cell grows, the midcell repressor concentration diminishes until Cdr2 (and thus the Cdk1 pathway) is no longer repressed and entry into mitosis is facilitated.

could be left behind at the cell poles after cytokinesis is complete and the divisome is disassembled. These proteins could provide targets for polar-localized proteins in the daughter cells. Just such a scenario has been proposed for TipN in *Caulobacter*, a landmark protein that provides a positional cue for the assembly of a flagellum at the pole and for controlling the size asymmetry of the two daughter cells (*48*, *49*). Hemispherical curvature and perhaps differences in cell envelope molecular composition are also potential ultimate cues for placing proteins at the cell poles. Conceivably, some proteins, including DivIVA and TipN, rely on both cues in their localization.

Understanding dynamic changes in protein localization. Finally, we touch on an important challenge for the future. So far we have focused on how proteins come to localize at a particular site, but not on the events that precipitate changes in protein localization. For example, what changes in the cell so that specific histidine kinases transiently localize to a cell pole rather than remaining dispersed throughout the cell membrane? What signals the replacement of that kinase with a different polar kinase? A contributing factor is changes in the phosphorylation state of one or more proteins that affects their collective binding affinity for the localization site. In other cases, the precipitating event is the de novo synthesis of a protein that then localizes by diffusion-and-capture at a target site. In each case of localized proteins involved in progression of the cell cycle or other serial developmental processes, the precipitation of localization is the consequence of an upstream event, such as a change in the cell's topology, the prior positioning of a target binding factor, a transcriptional cascade, a series of phospho-transfer reactions, and so on. These are all causal events involving physical processes, such as diffusion and stereochemistry, that are just beginning to be understood. The challenge now is to assemble the emerging evidence into a systems-level understanding of protein localization within the overall operation of the cell.

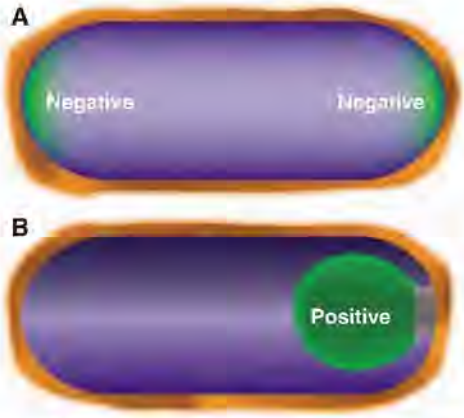


Fig. 4. Geometric cues for protein localization. **(A)** Schematic depiction of the hemispherical poles of the cell. The more extreme negative curvature at the inside surface of the poles (relative to the inside surface of the lateral walls of the cell) could be a geometric cue for proteins, such as the cell division protein DivIVA, that localize to the poles. **(B)** The process of engulfment during spore formation in *B. subtilis* in which the membrane of the large mother cell (on the left) migrates around to surround and eventually pinch off the nascent spore. The positive curvature of the surface of the engulfment membrane from within the mother cell is a geometric cue for the sporulation protein SpoVM. The green color indicates the localization of DivIVA in (A) and SpoVM in (B) to regions of negative and positive curvature, respectively.

Conclusions

Bacteria have an intricate and dynamic three-dimensional organization that is central to their capacity to grow and divide, to respond to the environment, and to develop into specialized cells. This high state of organization is not limited to proteins; the chromosome too is maintained in a strikingly organized manner. Why have bacteria evolved this intricate architecture, and how is it achieved? As we have seen, the inner life of the cell is inextricably linked to how the cell works. Proteins assemble at specific sites (such as the poles) to amplify signals from the environment, to prevent inappropriate cytokinesis, to capture chromosomal regions, to control proteolysis, and to dictate cell shape. But what are the ultimate cues that dictate position in the cell? The answers seem to lie in the interactions of certain proteins with each other through processes of dynamic self-assembly and in the very geometry of the cell itself. Clearly, a complete picture of the cell requires us to continually ask where proteins are, why they localize where they do, and how this localization is achieved.

References and Notes

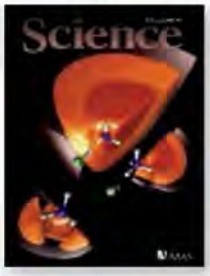
1. J. N. Werner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7858 (2009).
2. M. R. Alley, J. R. Maddock, L. Shapiro, *Genes Dev.* **6**, 825 (1992).
3. J. R. Maddock, L. Shapiro, *Science* **259**, 1717 (1993).
4. P. Zhang, C. M. Khursigara, L. M. Hartnell, S. Subramaniam, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3777 (2007).
5. A. Briegel *et al.*, *Mol. Microbiol.* **69**, 30 (2008).
6. T. S. Shimizu *et al.*, *Nat. Cell Biol.* **2**, 792 (2000).
7. P. Ames, C. A. Studdert, R. H. Reiser, J. S. Parkinson, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7060 (2002).
8. K. K. Kim, H. Yokota, S. H. Kim, *Nature* **400**, 787 (1999).
9. P. Ames, J. S. Parkinson, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9292 (2006).
10. N. Ausmees, J. R. Kuhn, C. Jacobs-Wagner, *Cell* **115**, 705 (2003).
11. R. B. Jensen, L. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10661 (1999).
12. C. D. Webb *et al.*, *Cell* **88**, 667 (1997).
13. M. A. Fogel, M. K. Waldor, *Mol. Microbiol.* **55**, 125 (2005).
14. G. R. Bowman *et al.*, *Cell* **134**, 945 (2008).
15. G. Ebersbach, A. Briegel, G. J. Jensen, C. Jacobs-Wagner, *Cell* **134**, 956 (2008).
16. S. Ben-Yehuda, D. Z. Rudner, R. Losick, *Science* **299**, 532 (2003); published online 19 December 2002 (10.1126/science.1079914).
17. L. J. Wu, J. Errington, *Mol. Microbiol.* **49**, 1463 (2003).
18. P. H. Viollier *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9257 (2004).
19. A. A. Teleman, P. L. Graumann, D. C. Lin, A. D. Grossman, R. Losick, *Curr. Biol.* **8**, 1102 (1998).
20. L. J. Wu, J. Errington, *Science* **264**, 572 (1994).
21. H. Niki, Y. Yamaichi, S. Hiraga, *Genes Dev.* **14**, 212 (2000).
22. J. Lutkenhaus, *Annu. Rev. Biochem.* **76**, 539 (2007).
23. A. L. Marston, H. B. Thomaides, D. H. Edwards, M. E. Sharpe, J. Errington, *Genes Dev.* **12**, 3419 (1998).
24. M. Bramkamp *et al.*, *Mol. Microbiol.* **70**, 1556 (2008).
25. J. E. Patrick, D. B. Kearns, *Mol. Microbiol.* **70**, 1166 (2008).
26. Q. Sun, W. Margolin, *J. Bacteriol.* **186**, 3951 (2004).
27. T. G. Bernhardt, P. A. de Boer, *Mol. Cell* **18**, 555 (2005).
28. L. J. Wu *et al.*, *EMBO J.* **28**, 1940 (2009).
29. L. J. Wu, J. Errington, *Cell* **117**, 915 (2004).
30. M. Thanbichler, L. Shapiro, *Cell* **126**, 147 (2006).
31. J. B. Moseley, A. Mayeux, A. Paoletti, P. Nurse, *Nature* **459**, 857 (2009).
32. S. G. Martin, M. Berthelot-Grosjean, *Nature* **459**, 852 (2009).
33. P. T. McGrath, A. A. Iniesta, K. R. Ryan, L. Shapiro, H. H. McAdams, *Cell* **124**, 535 (2006).
34. J. Kain, G. G. He, R. Losick, *J. Bacteriol.* **190**, 6749 (2008).
35. I. J. Domian, K. C. Quon, L. Shapiro, *Cell* **90**, 415 (1997).
36. K. C. Quon, B. Yang, I. J. Domian, L. Shapiro, G. T. Marczynski, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 120 (1998).
37. U. Jenal, T. Fuchs, *EMBO J.* **17**, 5658 (1998).
38. A. A. Iniesta, P. T. McGrath, A. Reisenauer, H. H. McAdams, L. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10935 (2006).
39. A. Duerig *et al.*, *Genes Dev.* **23**, 93 (2009).
40. H. Kobayashi, N. J. De Nisco, P. Chien, L. A. Simmons, G. C. Walker, *Mol. Microbiol.* **73**, 586 (2009).
41. S. R. Lybarger, T. L. Johnson, M. D. Gray, A. E. Sikora, M. Sandkvist, *J. Bacteriol.* **191**, 3149 (2009).
42. J. Deich, E. M. Judd, H. H. McAdams, W. E. Moerner, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15921 (2004).
43. D. Z. Rudner, R. Losick, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8701 (2002).
44. D. Greenfield *et al.*, *PLoS Biol.* **7**, e1000137 (2009).
45. K. S. Ramamurthi, S. Lecuyer, H. A. Stone, R. Losick, *Science* **323**, 1354 (2009).
46. R. Lenarcic *et al.*, *EMBO J.* **28**, 2272 (2009).
47. K. S. Ramamurthi, R. Losick, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13541 (2009).
48. E. Huitema, S. Pritchard, D. Matteson, S. K. Radhakrishnan, P. H. Viollier, *Cell* **124**, 1025 (2006).
49. H. Lam, W. B. Schofield, C. Jacobs-Wagner, *Cell* **124**, 1011 (2006).
50. Supported by NIH grants GM51426 and GM325062 (L.S.), NIH grant GM07311 (H.H.M. and L.S.), U.S. Department of Energy Office of Science grant DE-FG02-05ER64136 (H.H.M. and L.S.), and NIH grant GM18568 (R.L.).

10.1126/science.1175685

The Evolution of Gift Giving

For them: *Science* and full AAAS benefits.
For you: a special gift rate and our new holiday shirt!

$99 Professional gift rate (reg. $146)
$50 Postdoc/Student gift rate (reg. $99/$75)



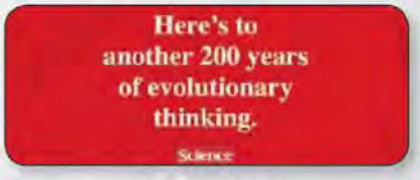
Make the holidays happier for a promising young researcher, family member, friend, or student. *Science* is the gift that lasts all year—and includes all the benefits of membership in AAAS.

Get a FREE limited-edition holiday shirt. With its playful image of Darwin and "¡Viva la Evolución!" attitude, the shirt has been a big hit in basic black. Now it can be yours in cheerful holiday red.

You'll enjoy something else as well: the satisfaction of helping to support AAAS and our international, public policy, and educational programs—the ones that advance science and serve society. Happy holidays, indeed!

Order now: visit promo.aaas.org/darwinholiday or call 866-434-2227.

A $22.50 value—yours FREE when you give *Science!*



Detail on back

Please order at least two weeks before the holiday you're celebrating, so we have time to send your gift recipient a letter announcing your gift. Non-U.S. recipients may receive *Science* Digital edition at the special gift rate. Check online for print edition rates. $74 allocated to *Science* for Professional memberships, $50 for Postdoc/Student memberships. Please allow 4 weeks for receipt of first issue. Prices valid through 1/31/10.

Science
AAAS

Induced Chromosomal Proximity and Gene Fusions in Prostate Cancer

Ram-Shankar Mani,[1,2] Scott A. Tomlins,[1,2] Kaitlin Callahan,[1,2] Aparna Ghosh,[1,2] Mukesh K. Nyati,[3,4] Sooryanarayana Varambally,[1,2,3] Nallasivam Palanisamy,[1,3] Arul M. Chinnaiyan[1,2,3,5,6]*

Gene fusions are a hallmark of cancer development (*1*), but the mechanisms underlying their genesis and cell type specificities are unclear (*2*).

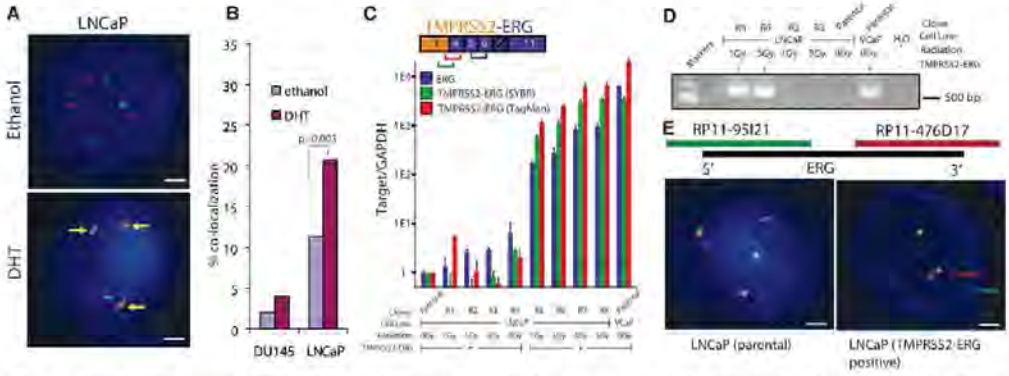


Fig. 1. (**A**) FISH-based evaluation of induced proximity between *TMPRSS2* (green) and *ERG* (red) on stimulation with 100 nM DHT in LNCaP cells. Colocalization is indicated by yellow arrows. Scale bars indicate 2 μm. (**B**) Induced proximity between *TMPRSS2* and *ERG* in DU145 and LNCaP cells is quantified and represented as the percentage of nuclei exhibiting colocalization signals. GAPDH, glyceraldehyde-3-phosphate dehydrogenase. (**C**) SYBR Green (Molecular Probes, Eugene, Oregon) and TaqMan (Applied Biosystems, Foster City, California) quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of the *TMPRSS2-ERG* fusion transcript. (**D**) Gel-based RT-PCR analysis with primers spanning the first exon of *TMPRSS2* and the sixth exon of *ERG* for representative clones. (**E**) FISH analysis of the *ERG* locus. Split signals representing *ERG* rearrangement are highlighted by arrows.

Fusions of the *TMPRSS2* and *ERG* genes, which are located 3 Mb apart on human chromosome 21q22.2, are found in about 50% of prostate cancers and consist of the 5′ untranslated region of *TMPRSS2*, an androgen-regulated gene, fused to the protein-coding sequences of *ERG*, which encodes an erythroblast transformation–specific (ETS) transcription factor (*2*). Recently, estrogen was shown to induce rapid chromosomal movements that bring together estrogen receptor α–bound genes on different chromosomes (*3*). Given the broad similarities between estrogen and androgen signaling, we hypothesized that androgen might likewise induce chromosomal movements and bring together the *TMPRSS2* and *ERG* genes, thereby facilitating the formation of gene fusion.

To examine the effects of androgen signaling on the formation of *TMPRSS2-ERG*, we studied LNCaP prostate cancer cells, which are androgen sensitive but lack this fusion gene (*4*) (fig. S1). We treated these cells with the androgen receptor (AR) ligand dihydrotestosterone (DHT) for 60 min. With use of dual-color fluorescence in situ hybridization (FISH), we observed that stimulation with DHT induced proximity between *TMPRSS2* and *ERG* genomic loci (34/300 and 62/300 nuclei for ethanol and DHT, respectively, $P = 0.003$, χ^2 test) (Fig. 1, A and B, and fig. S2). This effect was dependent on AR (fig. S3). Importantly, androgen did not induce proximity between the *TMPRSS2* and *ERG* loci in DU145 human prostate cancer cells, which are androgen insensitive (4/197 and 8/197 nuclei for ethanol and DHT, respectively, $P = 0.38$, χ^2 test).

To determine whether this induced proximity facilitates the formation of gene fusions, we treated LNCaP cells with DHT for 12 hours and then irradiated the cells to induce DNA double-strand breaks. After irradiation [1 or 3 grays (Gy)], single cells were seeded in multiple 96-well plates by using flow cytometry and were clonally expanded. *TMPRSS2-ERG* fusion transcripts were detected in 25% (3/12) of the clones treated with 3-Gy irradiation but in only 2.3% (1/43) of those treated with 1 Gy (Fig. 1, C and D, fig. S4, and table S1). The expression of *TMPRSS2-ERG* transcripts in positive LNCaP clones was similar to that in VCaP prostate cancer cells (*3*), which endogenously harbor the gene fusion. Furthermore, *TMPRSS2-ERG*–expressing LNCaP cells showed evidence consistent with chromosomal aberrations at the *ERG* locus (Fig. 1E and fig. S5). Parental LNCaP cells have four copies of chromosome 21, generating four yellow signals. LNCaP cells that harbor *TMPRSS2-ERG* transcripts generate three yellow signals and a pair of red and green signals, suggestive of a rearrangement at the *ERG* locus.

Our results build on earlier work documenting the role of chromosomal proximity in gene fusions (*5*) and, in so doing, provide a conceptual framework for the genesis of gene fusions in hormone-regulated epithelial cancers. Androgen-induced proximity between *TMPRSS2* and *ERG* could help explain why *TMPRSS2-ERG* fusions are restricted to the prostate, which is uniquely dependent on androgen signaling. We speculate that androgen signaling colocalizes the 5′ and 3′ gene fusion partners, thereby increasing the probability of a gene fusion when subjected to agents that cause DNA double-strand breaks.

References and Notes

1. J. D. Rowley, *Annu. Rev. Genet.* **32**, 495 (1998).
2. S. A. Tomlins *et al.*, *Science* **310**, 644 (2005).
3. Q. Hu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19199 (2008).
4. Materials and methods are available as supporting material on *Science* Online.
5. M. N. Nikiforova *et al.*, *Science* **290**, 138 (2000).
6. This work was supported by NIH grants P50CA69568, R01CA132874, and DOD W81XWH-09-2-0014. A.M.C. is a paid consultant to Gen-Probe, Incorporated. A.M.C. and S.A.T. are named as co-inventors on a patent related to gene fusions in prostate cancer that is licensed to Gen-Probe, Incorporated.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1178124/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References and Notes

23 June 2009; accepted 21 October 2009
Published online 29 October 2009;
10.1126/science.1178124
Include this information when citing this paper.

[1]Michigan Center for Translational Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA. [2]Department of Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA. [3]Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109, USA. [4]Department of Radiation Oncology, University of Michigan Medical School, Ann Arbor, MI 48109, USA. [5]Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, MI 48109, USA. [6]Department of Urology, University of Michigan Medical School, Ann Arbor, MI 48109, USA.

*To whom correspondence should be addressed. E-mail: arul@umich.edu

Haploid Genetic Screens in Human Cells Identify Host Factors Used by Pathogens

Jan E. Carette,[1] Carla P. Guimaraes,[1] Malini Varadarajan,[1] Annie S. Park,[1] Irene Wuethrich,[1] Alzbeta Godarova,[1] Maciej Kotecki,[2] Brent H. Cochran,[2] Eric Spooner,[1] Hidde L. Ploegh,[1,3] Thijn R. Brummelkamp[1]*

Loss-of-function genetic screens in model organisms have elucidated numerous biological processes, but the diploid genome of mammalian cells has precluded large-scale gene disruption. We used insertional mutagenesis to develop a screening method to generate null alleles in a human cell line haploid for all chromosomes except chromosome 8. Using this approach, we identified host factors essential for infection with influenza and genes encoding important elements of the biosynthetic pathway of diphthamide, which are required for the cytotoxic effects of diphtheria toxin and exotoxin A. We also identified genes needed for the action of cytolethal distending toxin, including a cell-surface protein that interacts with the toxin. This approach has both conceptual and practical parallels with genetic approaches in haploid yeast.

Identification of gene products that play a role in human disease drives much of today's biomedical research. Classical genetics with induced mutations, as pioneered by Muller in 1927 (*1*), is the most powerful unbiased approach to elucidate the genetic components that underlie biological processes. The study of cultured human cells allows the recapitulation of many essential elements of human disease. However, the inability to efficiently generate and recover bi-allelic mutants in human diploid cells limits the contribution of mutagenesis-based genetics to the understanding of human disease.

The identification of cellular genes exploited by bacteria and viruses is essential to elucidate the mechanisms by which these pathogens cause disease (*2*–*4*). Bacterial toxins contribute greatly to pathogenicity of the microbes that produce them. Identification of host proteins involved in toxin cytotoxicity should help to identify targets for therapeutic intervention in diseases caused by bacteria, many of which now show increased resistance to conventional antibiotics. Likewise, an understanding of how viruses depend on host proteins to enter the cell, replicate their genome, and spread may accelerate the development of antiviral drugs. Influenza virus remains a threat to human health, causing several hundred thousands of deaths annually and many more in the course of a pandemic (*5*). The rapid spread of new strains of influenza A [for instance, avian (H5N1) and swine (H1N1) influenza] and the emergence of drug-resistant influenza strains (*6*) limit the effectiveness of vaccines and current antiviral therapeutics. Thus, we developed a method for genetic screens in human cells and isolated genes required for the action of several bacterial toxins and influenza viruses.

Development of an approach for haploid genetic screens in human cells. To facilitate mutagenesis-based genetic approaches in human cells, we used a derivative of the 7 KBM7 chronic myeloid leukemia (CML) cell line with a haploid karyotype except for chromosome 8 (Fig. 1A) (*7*, *8*). In this cell line of hematopoietic origin, gene inactivation should allow the generation of null mutants for most nonessential genes. We chose to inactivate genes with the use of insertional mutagenesis because this approach is highly mutagenic in a variety of organisms, and the integrated DNA sequences provide a convenient molecular tag to identify the disrupted gene. We used gene-trap retroviruses that contain a strong adenoviral splice acceptor site and a marker gene (green fluorescent protein or puromycin-resistance gene) in reverse orientation of the retroviral backbone. To examine whether gene-trap insertions were indeed mutagenic, we performed a pilot screen to isolate mutants resistant to the nucleotide analog 6-thioguanine (6-TG), converted to a toxic metabolite by the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT). The gene-trap virus was titrated to obtain a single viral integration in the majority of the infected cells. Cell lines resistant to 6-TG were recovered, and five independent mutants carried insertions in intron 1 of the X-linked HPRT gene (Fig. 1B and fig. S1A). We next performed two genetic screens to target autosomal genes. KBM7 cells are sensitive to the tumor necrosis factor–related apoptosis-inducing ligand (TRAIL) and to inhibition of the BCR-ABL oncogenic fusion protein by the kinase inhibitor Gleevec (Novartis, Basel, Switzerland). Gene trap–mutagenized KBM7 cells were exposed to either TRAIL or Gleevec and

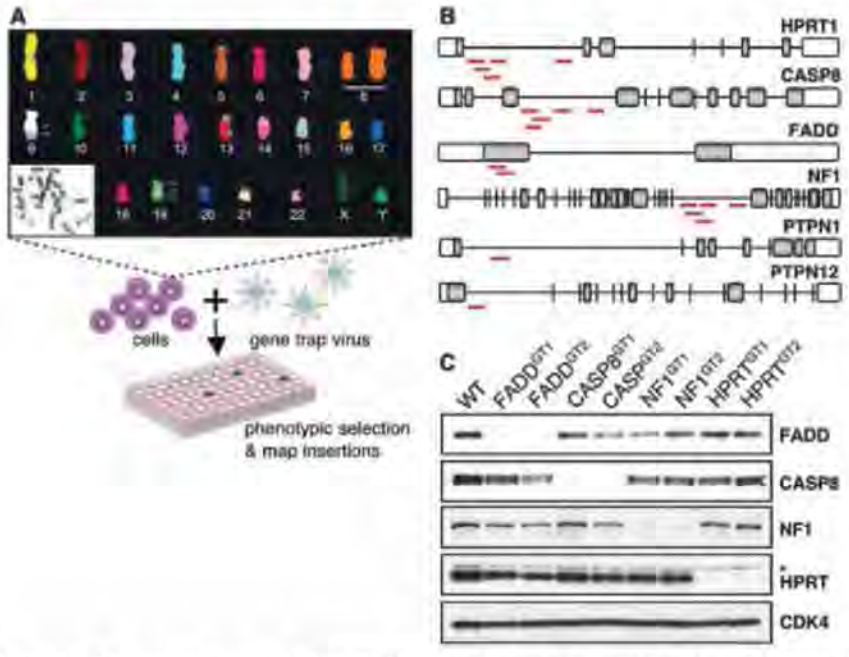


Fig. 1. Haploid genetic screens. (**A**) Twenty-four–color spectral karyotype of near-haploid KBM7 cells and schematic outline of gene trap–mutagenesis screens. (**B**) Schematic outline of the gene-trap insertion sites (red lines) in cells exposed to 6-TG, TRAIL, or Gleevec. (**C**) Immunoblot analysis of FADD, CASP8, NF1, and HPRT expression levels in clones that contain independent gene-trap insertions in the respective loci. Mutant alleles are labeled with GT in superscript notation, and an unspecific background band is indicated with an asterisk. CDK4 was used as a loading control.

[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [2]Department of Physiology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA.

*To whom correspondence should be addressed. E-mail: brummelkamp@wi.mit.edu

resistant mutants were recovered. Five TRAIL-resistant clones showed independent insertions in caspase 8 (CASP8) and two independent insertions in Fas-associated death domain protein (FADD), genes known to be required for TRAIL-induced apoptosis (Fig. 1B and fig. S1A) (*9*). Resistance to TRAIL was confirmed in these mutants (fig. S1B). Five independent Gleevec-resistant mutants contained insertions in neurofibromin 1 (NF1) and one in protein tyrosine phosphatase-N1 (PTPN1); both genes are known to play an important role in the response of CML cells to Gleevec (*10*). One insertion was found in PTPN12, a tyrosine phosphatase that interacts with c-abl and negatively regulates its activity (*11*). Thus, PTPN12 is critical for Gleevec sensitivity. All insertions were in the same transcriptional orientation as the target gene, and immunoblot analysis of HPRT, FADD, CASP8, and NF1 mutant cells failed to detect the corresponding gene products (Fig. 1C). The haploid background of KBM7 thus enables the generation of mutant alleles for autosomal genes and pinpoints genes involved in the biological processes under study.

Identification of host factors required for cytolethal distending toxin. Because many pathogenic agents such as bacterial toxins or viruses readily kill the cells they target, a large-scale production of knockout alleles for human genes may enable the identification of host factors essential for pathogenesis, such as enzymes that create structures recognized by toxins or viruses or the receptors themselves. Several pathogenic bacteria (such as *Escherichia coli*, *Shigella dysenteriae*, *Actinobacillus actinomycetemcomitans*, *Campylobacter jejuni*, *Helicobacter* spp., *Salmonella typhi*, and *Haemophilus ducreyi*) secrete potent bacterial toxins named cytolethal distending toxins (CDTs). The DNAse I–like CdtB subunit of these toxins enters the nucleus and causes cytotoxicity by inducing DNA breaks (*12*, *13*). The membrane receptor(s) and other essential host genes involved in the entry or action of CDTs are unknown. KBM7 cells responded to *E. coli*–derived CDT by undergoing an arrest in the G_2/M phase of the cell cycle (Fig. 2A) that precedes cell death. Mutagenized KBM7 cells were treated with CDT and resistant clones were isolated. Eleven independent insertions in sphingomyelin synthase 1 (SGMS1) and three insertions in TMEM181, a gene that encodes a putative G protein–coupled receptor (*14*), were recovered (Fig. 2B and fig. S2A). SGMS1 and TMEM181 mutants were resistant to CDT, a phenotype reverted by complementing the mutant cells with the corresponding cDNAs (Fig. 2C and fig. S2B). The SGMS1 mutation reduced levels of sphingomyelin, as verified by treatment of cells with lysenin, a sphingomyelin-specific pore-forming toxin (fig. S3A). Sphingomyelin is a key component of lipid rafts: Depletion of SGMS1 activity disturbs lipid-raft function and prevents receptor clustering (*15*), a trait of possible relevance for CDT binding and/or entry. Extraction of the lipid-raft component cholesterol from the plasma membrane abolishes CDT binding (*16*).

TMEM181 mutants remained fully sensitive to lysenin (fig. S3, A and B), suggesting that their resistance is acquired by means other than sphingomyelin depletion. Because TMEM181 is present at the cell surface (*17*) and a receptor for CDT must localize to the plasma membrane, we tested whether CDT bound to TMEM181. Flag-tagged CDT was adsorbed onto anti-Flag beads and incubated with cell lysates prepared from wild-type (WT) KBM7 cells and from KBM7 cells that express hemagglutinin (HA)–tagged TMEM181. Immunoblot analysis showed robust binding of TMEM181 to CDT (Fig. 2D). When TMEM181 was overexpressed by retroviral transduction in NIH3T3, U2OS, and Hela cells, it sensitized them to CDT intoxication (Fig. 2E and fig. S4), suggesting that TMEM181 expression levels are rate limiting for intoxication. Thus, CDT may bind to the cell-surface protein TMEM181, an event both required and rate limiting for intoxication, and then enter the cell through sphingomyelin-dependent, lipid-raft–mediated endocytosis, followed by nuclear entry and cleavage of cellular DNA. However, these results do not rule out the possibilities that TMEM181 is part of a complex that constitutes a functional receptor or plays a role in trafficking of a receptor-toxin complex.

Isolation of host factors essential for influenza virus infectivity. We next isolated mutant cells that were resistant to influenza virus A (PR/8/34; H1N1). Proviral-host junction sequencing revealed two independent insertions in cytidine monophosphate *N*-acetylneuraminic acid synthase (CMAS), encoding the enzyme responsible for activation of NeuAc to CMP-NeuAc, the glycosyl donor used in sialic acid–containing glycoconjugate synthesis. These structures are the receptors on influenza-susceptible cells recognized by the influenza hemagglutinin. We recovered three independent insertions in SLC35A2 (Fig. 3A and fig. S5A), a gene whose product transports uridine 5′-diphosphate–galactose from the cytoplasm to the Golgi, where it serves as a glycosyl donor (*18*) important for the generation of glycans to be modified with sialic acids. To determine whether mutant cells could be infected by influenza, we exposed cells to the virus and

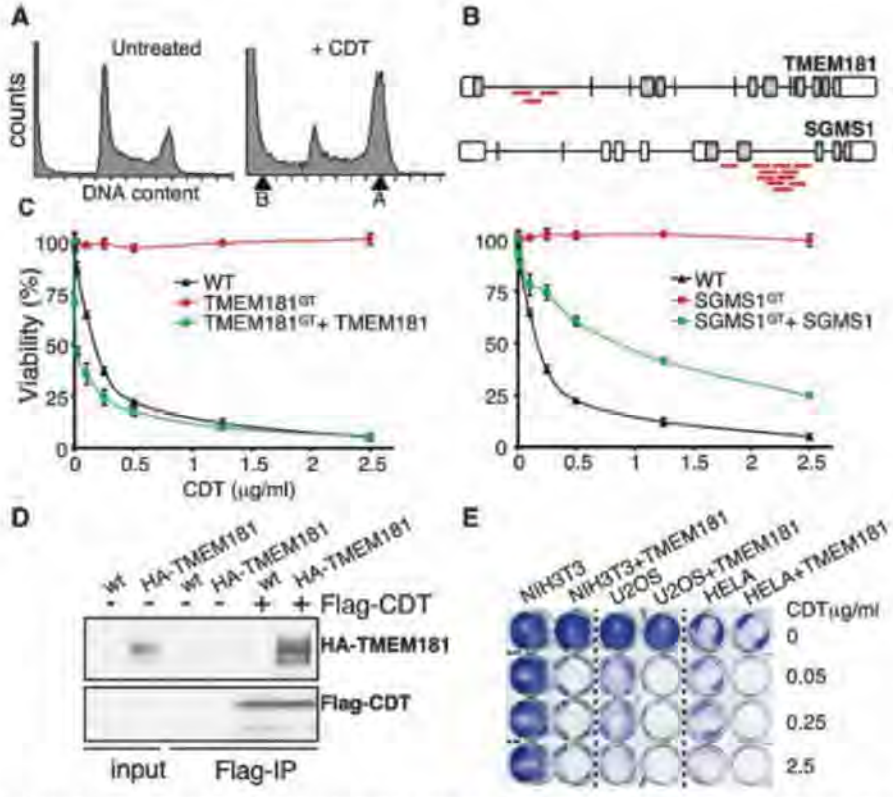


Fig. 2. Host factors for CDT. (**A**) Flow cytometric analysis of control KBM7 cells (left) and KBM7 cells after exposure to CDT purified from *E. coli* (right). Exposure of cells to CDT results in an increase of cells in the G_2/M phase of the cell cycle (arrow A) and cell death (arrow B). (**B**) Schematic outline of the insertion sites (red lines) in mutant cells unresponsive to CDT. (**C**) Resistance of TMEM181 mutant cells and SGMS1 mutant cells to CDT. Mutant cells reconstituted with the respective cDNAs re-acquire toxin sensitivity. Results are presented as mean values ± SD (error bars) ($n = 3$ biological replicates). (**D**) Flag-tagged CDT was bound to immobilized anti-Flag antibody and incubated with KBM7 cell lysates or lysates of cells expressing HA-TMEM181. Bound proteins were eluted and subjected to immunoblot analysis. Anti-Flag beads without bound CDT served as a control. (**E**) NIH3T3, U2OS, and HELA cells infected with a TMEM181-expressing retrovirus were treated with increasing amounts of CDT. After 5 days, adherent cells were stained with crystal violet.

stained for influenza nucleoprotein 12 hours after infection. As expected, KBM7 cells showed high levels of infection (~95% infection), whereas CMAS and SLC35A2 mutant cells showed near-complete resistance to virus infection (<0.01% infection) (Fig. 3B and fig. S5D). Absence of CMAS and SLC35A2 expression in the mutants was verified by reverse transcription polymerase chain reaction (RT-PCR) or immunoblot analysis (fig. S5, B and C). Transduction with cDNAs encoding the disrupted genes fully restored susceptibility to influenza infection (Fig. 3B), indicating that the observed resistance is attributable to the mutated loci. Although the KBM7 genome has not been screened at saturation for resistance to influenza, the transporter (SLC35A2) and enzyme (CMAS) identified here could lead to the development of antiviral therapies for influenza.

Identification of host factors for ADP-ribosylating toxins. Diphtheria and anthrax toxins are AB toxins composed of a cell-binding moiety (B) and an active (A) subunit that targets a host function to increase virulence. We have a detailed molecular understanding of how diphtheria toxin enters the cell and induces cell death (*19*, *20*). Can haploid genetic screens identify previously unidentified components in this well-characterized host/pathogen interaction? We screened mutagenized cells with diphtheria or anthrax toxin. Because native anthrax toxin is not cytotoxic for KBM7 cells, we exposed cells to the cell-binding component of anthrax toxin–protective antigen (PA) and anthrax lethal factor (LFN) fused to the catalytic domain of diphtheria toxin (LFN-DTA) (*21*). Resistant mutants were classified as either being resistant to anthrax toxin (PA-LFN-DTA) (class I), resistant to diphtheria toxin (class II), or resistant to both (class III). Mutants in the known anthrax toxin receptor (ANTXR2) (*22*) were recovered with 10 independent insertions and with 12 insertions for the known diphtheria toxin receptor [heparin-binding EGF-like growth factor (HB-EGF)] (Fig. 4A and fig. S6A) (*23*). The third class of mutants included genes involved in diphthamide biosynthesis [DPH1, DPH2, and DPH5; see (*20*)] and a previously uncharacterized gene named WDR85 (Fig. 4A and fig. 6A). All of these insertions were in the same transcriptional orientation as the mutated gene and were thus predicted to impair gene function. In the WDR85 mutant (hereafter referred to as $WDR85^{GT}$), no WDR85 transcripts were observed, as determined by RT-PCR (Fig. 4B). The resistance of $WDR85^{GT}$ was readily complemented by transduction with WDR85 cDNA, which restored the sensitivity of $WDR85^{GT}$ cells to diphtheria toxin and anthrax toxin (PA-LFN-DTA) (Fig. 4C and fig. S6B). $WDR85^{GT}$ cells were also resistant to *Pseudomonas* exotoxin A, another adenosine diphosphate (ADP)–ribosylating toxin with a similar mode of action as diphtheria toxin (Fig. 4D). Although native anthrax toxin is not lethal to most cell types, including KBM7, its cellular entry and activity can be probed by monitoring cleavage of its cellular target MEK-3. $WDR85^{GT}$ cells were still responsive to the native anthrax toxin, because the extent of proteolytic cleavage of MEK-3 was comparable for $WDR85^{GT}$ and WT cells (fig. S7A), suggesting that toxin entry was normal in $WDR85^{GT}$ cells.

WDR85 is part of the dipthamide biosynthetic pathway. Given the resistance of WDR85 mutant cells to different bacterial toxins, we further explored the mechanism by which WDR85 conferred sensitivity to toxin-mediated cell death. Diphtheria toxin, LFN-DTA, and exotoxin A inhibit host translation through ADP ribosylation of elongation factor 2 (EF2) and thus cause cell death (*19*). ADP ribosylation occurs on diphthamide, a posttranslationally modified histidine uniquely present in EF2 and conserved among all eukaryotes. As WDR85 was not required for toxin entry, we investigated EF2 ribosylation in response to diphtheria toxin. In cell lysates derived from $WDR85^{GT}$ cells, EF2 ADP ribosylation was impaired and could be restored by re-expression of a WDR85 cDNA (fig. S7B). EF2 fused to a streptavidin-binding peptide (SBP) purified from $WDR85^{GT}$ cells was also a poor substrate for ADP ribosylation in vitro (Fig. 5A). Impaired ADP ribosylation is thus an inherent property of EF2 derived from $WDR85^{GT}$ cells and is not due to the presence or absence of other factors present in cell lysates. Diphthamide biosynthesis is the result of stepwise posttranslational modification of His^{175} (fig. S7D), the proteins responsible for which are known (*20*, *24*, *25*). The second step involves the trimethylation of "intermediate" EF2 by the methyltransferase DPH5, with *S*-adenosylmethionine as the methyl donor (*26*). To investigate if this methylation step was affected by the loss of WDR85, we purified intermediate EF2 from DPH5 null cells and performed in vitro methylation assays in cell lysates. Efficient methylation of intermediate EF2 by WT and $WDR85^{GT}$ cell lysates suggested that WDR85 is not required for the second step of diphthamide biosynthesis (Fig. 5B). Next, we purified EF2 from $WDR85^{GT}$ cells and used LC/MS/MS (liquid chromatography–tandem mass spectrometry) to monitor the relevant modifications of His^{175}. Modifications of His^{175} predict an increase in mass by +143 (diphthamide), +142 (diphthine), and +101 (the intermediate) mass units for those peptides that carry the modified His residue. SBP-tagged EF2 isolated from $WDR85^{GT}$ showed a mass consistent with the presence of unmodified His^{175} (Fig. 5C), whereas modifications of EF2 purified from WT and DPH5 mutant cells showed a mass that was expected for the presence of diphthamide and intermediate, respectively (fig. S8, A and B). The absence of modified histidine in EF2 suggests that WDR85 plays a role in the first step in diphthamide biosynthesis.

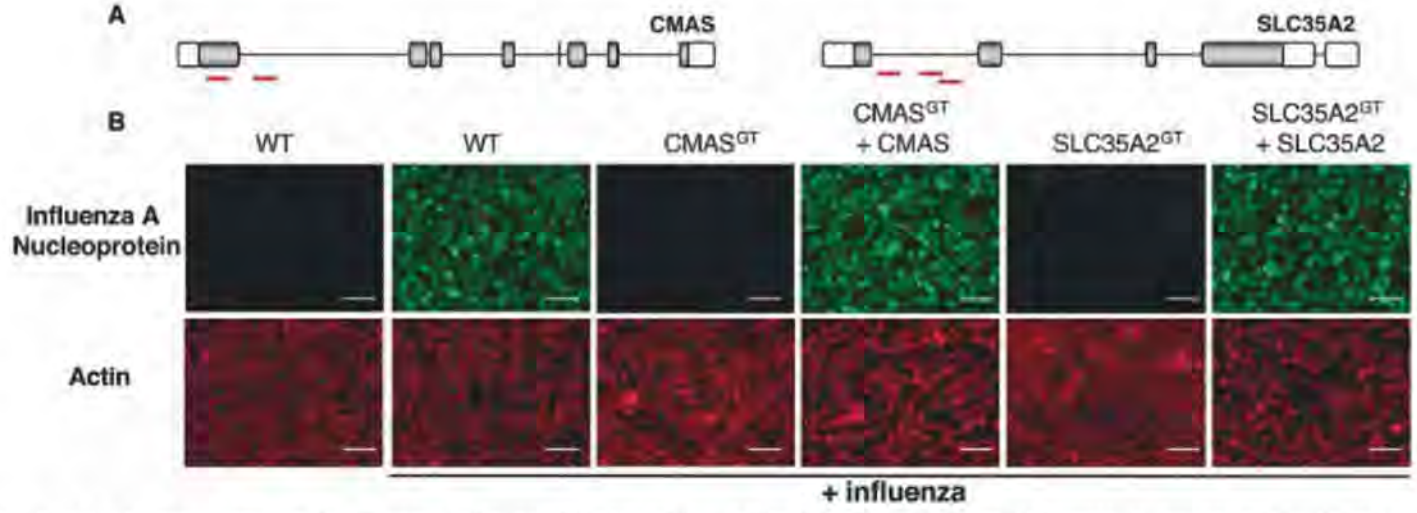


Fig. 3. Cellular genes required for influenza infection. (**A**) Schematic outline of the identified insertion sites (red lines) in the mutated genes. (**B**) Cells were exposed to influenza virus and stained 12 hours later with antibodies directed against influenza A nucleoprotein. Mutant cells reconstituted with cDNAs that correspond to the mutated gene products re-acquire virus sensitivity. Scale bars, 50 μm.

In the course of purification of EF2 from WDR85^{GT} cells, we detected a protein that strongly interacted with EF2 (Fig. 5D). Mass spectrometry and immunoblot analysis identified this protein as DPH5 (Fig. 5E and fig. S7C). WDR85 lacks homology to known proteins involved in diphthamide biosynthesis but does contain WD40 repeats, often involved in protein/protein interactions. Thus, WDR85 may serve as a scaffold to coordinate the association (or dissociation) of enzymatic complexes required for the stepwise biosynthesis of diphthamide.

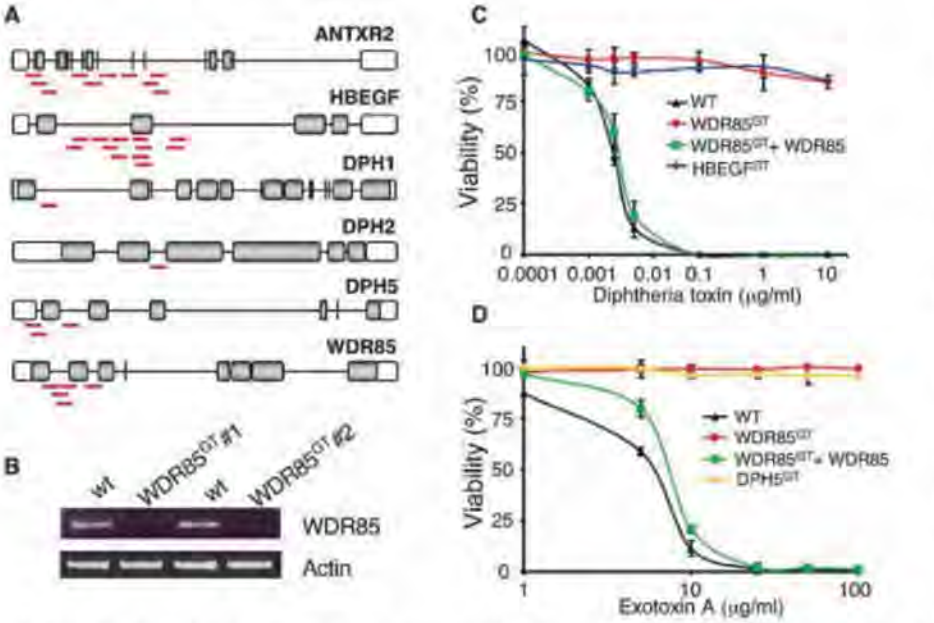


Fig. 4. Identification of loci that confer resistance to ADP-ribosylating bacterial toxins. (**A**) Schematic outline of the insertion sites (red lines) in the mutated genes. (**B**) RT-PCR for WDR85 shows undetectable WDR85 mRNA levels in independent clones with gene-trap insertions in the WDR85 locus. (**C**) Resistance of WDR85^{GT} cells to diphtheria toxin. Error bars indicate SD. (**D**) Resistance of WDR85^{GT} cells to exotoxin A. Identified clones with mutations in HB-EGF and DPH5 served as insensitive controls for these respective toxins, and WDR85^{GT} cells reconstituted with a WDR85 cDNA (WDR85^{GT} + WDR85) re-acquired sensitivity to the toxins. Results are presented as mean values ± SD (n = 3).

WDR85 is a conserved protein with homology to yeast YBR246W (fig. S9A). We used a database containing fitness profiles of deletion strains of all nonessential yeast genes under 1144 chemical conditions to cluster genes with similar profiles to YBR246W (*27*). The top 10 genes that phenoclustered with YBR246W by homozygous cosensitivity included DPH2 and DPH5 (fig. S9, B and C). The only gene annotation terms we found enriched concerned diphthamide biosynthesis [P value = 9×10^{-4} (fig. S9C)]. To directly test whether YBR246W is involved in diphthamide biosynthesis, we undertook ribosylation assays in protein extracts derived from WT yeast or yeast strains deleted for YKL191W (DPH2) or YBR246W. Deficiency of YKL191W and YBR246W both impaired ADP-ribosyl acceptor activity of EF2 in yeast (Fig. 5F). Thus, the role of WDR85 in diphthamide biosynthesis is conserved in eukaryotes, and the proposed scaffolding role may be the main function of WDR85 in cells. WDR85 therefore represents a host gene involved in the first step in diphthamide biosynthesis, despite previous suggestions that all proteins involved in this complex posttranslational modification were known (*20*).

Perspective. One of the main strengths of yeast as a genetic tool is the relative ease with which recessive mutations can be recognized and characterized at its haploid life stage (*28*), a trait absent from the commonly studied higher eukaryotes. The approach described here will allow similar types of screens for the human genome.

The discovery of RNA interference has enabled targeted reduction of gene expression in diploid cultured mammalian and insect cells, which

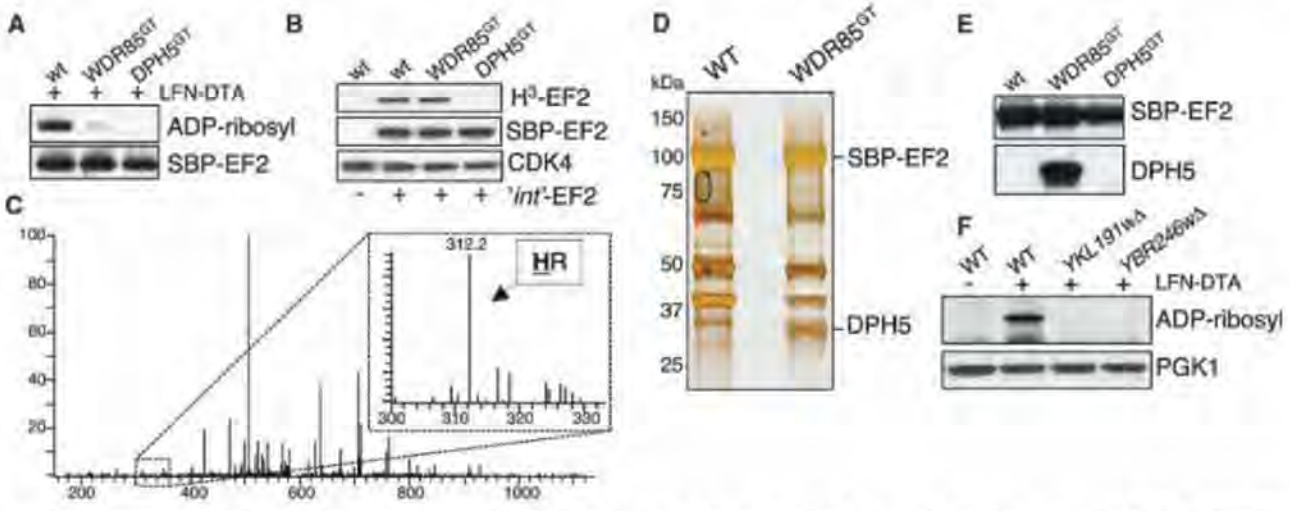


Fig. 5. WDR85 is involved in diphthamide biosynthesis. (**A**) In vitro ADP ribosylation of SBP-tagged EF2 purified from WT, WDR85, and DPH5 mutant cells by LFN-DTA in the presence of nicotinamide adenine dinucleotide (NAD)–biotin. Streptavidin–horseradish peroxidase (HRP) was used to detect ADP ribosylation, and total EF2 was detected by immunoblot analysis. (**B**) Methylation of intermediate EF2 by WT, WDR85, and DPH5 mutant cell lysates. SBP-tagged intermediate EF2 was purified from DPH5 mutant cells and incubated in lysates derived from the indicated genotypes in the presence of [methyl-^{3}H] adenosylmethionine (Ado-S-Me) as methyl donor. The amount of supplied intermediate EF2 was detected by immunoblot analysis, with CDK4 as a loading control. (**C**) MS/MS spectra of a tryptic peptide derived from SBP-tagged EF2 purified from WDR85 mutant cells. Peptide fragments characteristic for unmodified His^{715} are indicated. (**D**) Silver stain of SPB-EF2 purified from WT and WDR85-deficient cells. kDa, kilodaltons. (**E**) Immunoblot analysis of SBP-EF2 pull-down from WT, WDR85-, and DPH5-deficient cells probed with an antibody directed against DPH5. (**F**) Protein extracts from WT, YKL191W-, and YBR246W-deficient *S. cerevisiae* strains were incubated with LFN-DTA in the presence of NAD-biotin. Streptavidin-HRP was used to detect ADP ribosylation, and PGK1 was used as loading control.

opened the door to large-scale screens. At the same time, limitations of this approach are increasingly apparent, such as the induction of off-target effects that complicate genome-wide screens in particular (*29*, *30*) and the inability to completely switch off gene expression. When similar small interfering RNA screens are conducted independently in mammalian cells, the lack of concordance between them is an additional complicating factor (*31*, *32*). Finally, mammals are rather robust in their tolerance to partial loss of gene function: Haploinsufficiency appears to be the exception rather than the rule, because inactivation of one gene copy, as in heterozygous knockout mice, rarely leads to severe phenotypes.

Although we have focused on host-pathogen biology, similar screens could in principle be applied to any phenotype that can be recognized in a population of mutant cells, such as modulation of a genetically encoded reporter. In the future, haploid genetic screens could be used to generate comprehensive compendia of host factors that are used by different pathogens and may yield new strategies to combat infectious disease. In conclusion, the haploid genetic screens described here expand mutagenesis-based screens in model organisms by providing a window on disease-associated molecular networks that can be studied in cultured human cells.

References and Notes

1. H. J. Muller, *Science* **66**, 84 (1927).
2. A. L. Brass *et al.*, *Science* **319**, 921 (2008); published online 10 January 2008 (10.1126/science.1152725).
3. J. A. Philips, E. J. Rubin, N. Perrimon, *Science* **309**, 1251 (2005); published online 14 July 2005 (10.1126/science.1116006).
4. L. Hao *et al.*, *Nature* **454**, 890 (2008).
5. R. Salomon, R. G. Webster, *Cell* **136**, 402 (2009).
6. A. Moscona, *N. Engl. J. Med.* **360**, 953 (2009).
7. M. Kotecki, P. S. Reddy, B. H. Cochran, *Exp. Cell Res.* **252**, 273 (1999).
8. Materials and methods are available as supporting material on *Science* Online.
9. S. Nagata, *Cell* **88**, 355 (1997).
10. B. Luo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20380 (2008).
11. F. Cong *et al.*, *Mol. Cell* **6**, 1413 (2000).
12. M. Lara-Tejero, J. E. Galán, *Science* **290**, 354 (2000).
13. D. Nesic, Y. Hsu, C. E. Stebbins, *Nature* **429**, 429 (2004).
14. M. Wistrand, L. Kall, E. L. L. Sonnhammer, *Protein Sci.* **15**, 509 (2006).
15. M. Miyaji *et al.*, *J. Exp. Med.* **202**, 249 (2005).
16. L. Guerra *et al.*, *Cell. Microbiol.* **7**, 921 (2005).
17. B. Wollscheid *et al.*, *Nat. Biotechnol.* **27**, 378 (2009).
18. H. Sprong *et al.*, *Mol. Biol. Cell* **14**, 3482 (2003).
19. R. J. Collier, *Toxicon* **39**, 1793 (2001).
20. S. Liu, G. T. Milne, J. G. Kuremsky, G. R. Fink, S. H. Leppla, *Mol. Cell. Biol.* **24**, 9487 (2004).
21. J. C. Milne, S. R. Blanke, P. C. Hanna, R. J. Collier, *Mol. Microbiol.* **15**, 661 (1995).
22. H. M. Scobie, G. J. A. Rainey, K. A. Bradley, J. A. T. Young, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5170 (2003).
23. J. G. Naglich, J. E. Metherall, D. W. Russell, L. Eidels, *Cell* **69**, 1051 (1992).
24. L. C. Mattheakis, W. H. Shen, R. J. Collier, *Mol. Cell. Biol.* **12**, 4026 (1992).
25. S. Liu, S. H. Leppla, *Mol. Cell* **12**, 603 (2003).
26. J. Y. Chen, J. W. Bodley, *J. Biol. Chem.* **263**, 11692 (1988).
27. M. E. Hillenmeyer *et al.*, *Science* **320**, 362 (2008).
28. S. L. Forsburg, *Nat. Rev. Genet.* **2**, 659 (2001).
29. Y. Ma, A. Creanga, L. Lum, P. A. Beachy, *Nature* **443**, 359 (2006).
30. C. J. Echeverri *et al.*, *Nat. Methods* **3**, 777 (2006).
31. S. P. Goff, *Cell* **135**, 417 (2008).
32. F. D. Bushman *et al.*, *PLoS Pathog.* **5**, e1000437 (2009).

33. We thank D. Sabatini, S. Nijman, J. Roix, and J. Pruszak for discussion and critical review of the manuscript; C. Y. Wu and G. Fink for yeast deletion strains; J. Kaper for the CDT expression plasmid; J. Collier, R. Moon, and M. Wernig for plasmids; and E. Guillen for help with influenza infections. C.P.G. has a fellowship from Fundacao Ciencia Tecnologia, Portugal. T.R.B. was funded by the Kimmel Foundation and the Whitehead Institute Fellows Program. The Whitehead Institute has filed a patent on the application of gene-trap mutagenesis in haploid or near-haploid cells to identify human genes that affect cell phenotypes, including host factors used by pathogens.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1231/DC1
Materials and Methods
Figs. S1 to S9
References

10 July 2009; accepted 5 October 2009
10.1126/science.1178955

Proteome Organization in a Genome-Reduced Bacterium

Sebastian Kühner,[1]* Vera van Noort,[1]* Matthew J. Betts,[1] Alejandra Leo-Macias,[1] Claire Batisse,[1] Michaela Rode,[1] Takuji Yamada,[1] Tobias Maier,[2] Samuel Bader,[1] Pedro Beltran-Alvarez,[1] Daniel Castaño-Diez,[1] Wei-Hua Chen,[1] Damien Devos,[1] Marc Güell,[2] Tomas Norambuena,[3] Ines Racke,[1] Vladimir Rybin,[1] Alexander Schmidt,[4] Eva Yus,[2] Ruedi Aebersold,[4] Richard Herrmann,[5] Bettina Böttcher,[1]† Achilleas S. Frangakis,[1] Robert B. Russell,[1] Luis Serrano,[2,6] Peer Bork,[1]‡ Anne-Claude Gavin[1]‡

The genome of *Mycoplasma pneumoniae* is among the smallest found in self-replicating organisms. To study the basic principles of bacterial proteome organization, we used tandem affinity purification–mass spectrometry (TAP-MS) in a proteome-wide screen. The analysis revealed 62 homomultimeric and 116 heteromultimeric soluble protein complexes, of which the majority are novel. About a third of the heteromultimeric complexes show higher levels of proteome organization, including assembly into larger, multiprotein complex entities, suggesting sequential steps in biological processes, and extensive sharing of components, implying protein multifunctionality. Incorporation of structural models for 484 proteins, single-particle electron microscopy, and cellular electron tomograms provided supporting structural details for this proteome organization. The data set provides a blueprint of the minimal cellular machinery required for life.

Biological function arises in part from the concerted actions of interacting proteins that assemble into protein complexes and networks. Protein complexes are the first level of cellular proteome organization: functional and structural units, often termed molecular machines, that participate in all major cellular processes. Complexes are also highly dynamic in the sense that their organization and composition vary in time and space (*1*), and they interact to form higher level networks; this property is central to whole-cell functioning. However, general rules concerning protein complex assembly and dynamics remain elusive.

The combination of affinity purification with mass spectrometry (MS) (*2*) has been applied to several organisms to provide a growing repertoire of molecular machines. Genome-wide screens in *Saccharomyces cerevisiae* (*3*–*5*) captured discrete, dynamic proteome organization and revealed higher-order assemblies with direct connections between complexes and frequent sharing of common components. To date these exhaustive analyses have been applied only in yeast. In bacteria, genome-wide yeast two-hybrid analyses have been reported (*6*, *7*), but only a few biochemical analyses on selected sets of complexes are available (*8*–*11*). The understanding of proteome organization in these organisms concerns thus the binary interaction networks.

Here, we report a genome-scale analysis of protein complexes in the bacterium *Mycoplasma pneumoniae*, a human pathogen that causes atypical

[1]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [2]Centro Regulacion Genomica–Universidad Pompeu Fabra, Dr Aiguader 88, 08003 Barcelona, Spain. [3]Pontificia Universidad Catolica de Chile, Alameda 340, Santiago, Chile. [4]ETH (Eidgenössische Technische Hochschule) Zürich, Wolfgang-Pauli-Strasse 16, 8093 Zürich, Switzerland; Faculty of Science, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland, and Institute for Systems Biology, Seattle, WA 98013, USA. [5]ZMBH (Zentrum für Molekulare Biologie der Universität Heidelberg), Im Neuenheimer Feld 282, 69120 Heidelberg, Germany. [6]ICREA (Institució Catalana de Recerca i Estudis Avançats), 08010 Barcelona, Spain.

*These authors contributed equally to this work.
†Present address: University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JR.
‡To whom correspondence should be addressed. E-mail: gavin@embl.de (A.-C.G.); bork@embl.de (P.B.)

pneumonia (*12*). This self-replicating organism has one of the smallest known genomes (689 protein-encoding genes) (*13*, *14*), making it an ideal model organism for the investigation of absolute essentiality (*15*). This analysis and the integration with other consistently derived large-scale data sets provide a blueprint of the proteome organization in a minimal cell and reveal principles underlying adaptation to a reduced genome.

Genome-wide screen for protein complexes in *M. pneumoniae*. We adapted the tandem affinity purification–mass spectrometry (TAP-MS) protocol (*2*) to *M. pneumoniae* M129 (Fig. 1) (*16*). We processed all 689 *M. pneumoniae* protein-coding genes, of which 617 were successfully cloned [90% of the genome (*14*)]. With use of a transposon-based expression system, we constructed a total of 456 *M. pneumoniae* strains. They carry a stable genomic integration of carboxy-terminal TAP fusions under transcriptional control of the *M. pneumoniae* clpB (mpn531) promoter. From this collection, all 352 individual strains expressing soluble TAP fusions were grown to confluence in 2 liters of adherent culture, leading to 212 successful purifications. The components of the purified complex were separated by denaturing gel electrophoresis, and individual bands were trypsin-digested and analyzed by MS (table S1). We processed a total of 10,447 MS samples and identified proteins by using a new approach that integrates the Mascot (*17*) and Aldente (*18*) search algorithms (*19*). This increased the identification of known complex components by ~20% compared with either method alone (fig. S1, A and B). The procedure also scores the quality of individual identifications by considering all peptide profiles that we observed for each protein, including our purification data set and a PeptideAtlas, a comprehensive set of tryptic peptides (*20*) measured with Fourier transform–MS from whole *M. pneumoniae* lysates (table S2). We removed protein identifications with overlapping peptide profiles (3%) (fig. S1, C and D). When applied to the entire purification data set, this approach uncovered 411 distinct proteins from 5899 identifications (table S2).

The 411 proteins identified with 212 tagged proteins correspond to 60% of the annotated open reading frames (ORFs) and 85% of the predicted soluble proteome (fig. S2). They cover all cellular functions, although low abundant, small, or trans-membrane segment–containing proteins are notably underrepresented (fig. S2). Membrane proteins purification requires separate biochemical protocols, so they were not included in this screen. The proportion of new proteins identified per purification dropped asymptotically as the screen progressed, implying that the procedure was near saturation (fig. S3). This entails recurring protein complex retrieval through reverse tagging and is important both to confirm novel interactions and to identify dynamic complexes (*3*).

To define complexes in a quantitative way, we first calculated socio-affinity indices that measure the frequency with which pairs of proteins were found associated in our set of biochemical purifications (*3*, *16*). We improved the concept by integrating predicted interactions from the STRING database (*21*) and the relative abundance of a given prey when associated with different baits (i.e., across different purifications) (*22*). We used the MS scores that measure the probability for a peptide mass fingerprint to characterize each protein based on spectral counting. A reduced score for a prey in a purification, when compared to the same prey in other purifications, reflects identifications by a smaller number of peptides (lower spectral counts); it is indicative of a spurious interaction and is therefore down-weighted (fig. S4A). We applied this new scoring scheme to the entire data set and calculated a list of 10,083 interactions. A cut-off was defined at an accuracy, that is, a fraction of true interactions (*23*), of more than 80%, which gave a set of 1058 high-confidence interactions (fig. S4, B and C; also table S3). We also measured the overall experimental reproducibility on a set of 18 experiments that we performed twice; duplicates included growth of adherent cultures, biochemical purifications, and MS analyses (*16*). For protein pairs with socio-affinity scores ≥0.8, the overall reproducibility is 73%; for those scoring below it is 43% ($P = 10^{-13}$, χ^2 test). For comparison, the reproducibility calculated on the duplicated MS measurements of 72 MS samples is 97%. We then applied cluster analysis by using a procedure called clique percolation that allows proteins to be part of different complexes. We varied the clustering parameters over reasonable ranges. The best conditions in terms of coverage (see below) generated a collection of 116 heteromultimeric complexes. They are organized into densely (>one link) and loosely interconnected (one link) components we called "core" and "attachment," respectively (fig. S4D and table S4). Generally, *M. pneumoniae* proteins within complexes and cores are more often co-expressed (*24*) and conserved between species than average; proteins within complexes appear on average in 244 species compared with 173 for the entire proteome (median = 190). Comparison to a set of 31 known complexes, described in other species (table S5), revealed a coverage

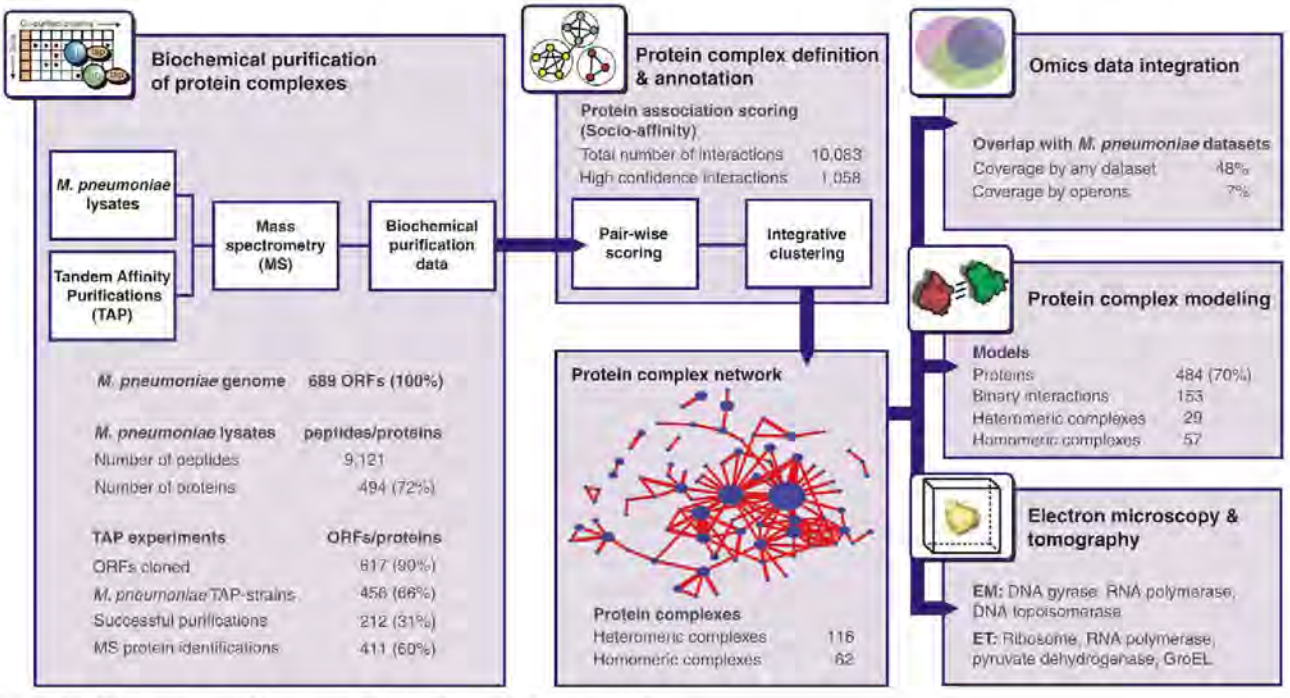


Fig. 1. Synopsis of the genome-wide screen of complexes in *M. pneumoniae*.

of 61%, which is similar to results from previous screens in yeast and *Escherichia coli* (coverage ~60%) (*3*, *4*, *9*, *25*).

Systematic detection of homomultimeric protein complexes. The TAP fusions were expressed from exogenous loci and promoter and are therefore present together with the untagged wild-type allele. It was thus common to observe both TAP-tagged and -untagged versions of the bait in the same purification, which is an indication of homomultimerization (fig. S5). Careful scrutiny of the purification data set revealed evidence for 62 homomultimeric complexes (table S4) covering 62% of those previously seen either in *M. pneumoniae* or in another species by orthology (table S5). Fourteen homomultimeric complexes were previously unknown, and for 12 of these we could find supporting structural evidence from homologs of known structure (*26*) (table S6). An example is Mpn266, a protein of previously unknown function that we found associated to RNA polymerase (complex 49, table S4) as a dimer. Its binding to the polymerase is consistent with its similarity to SpxA, an RNA polymerase–binding protein that regulates transcription initiation in Gram-positive bacteria (*27*, *28*). Comparative modeling of structure and single-particle electron microscopy (fig. S6) (*16*) show that *M. pneumoniae* RNA polymerase resembles that of *Thermus aquaticus* (*29*) with the exception of a substantially bigger stalk at the position of the sigma factor, RpoD (Mpn352), consistent with *M. pneumoniae* RpoD being 200 amino acids longer than its *T. aquaticus* ortholog (fig. S7A). The models also further support the idea that each Mpn266 in the dimer binds one of the two α subunits of the polymerase, as do other transcription factors (fig. S7A).

From the number of baits used (212) and from the effectiveness of the method in recovering known complexes (62% coverage), we estimate that as many as 47% of all soluble proteins form homomultimers in *M. pneumoniae*. This is in agreement with a recent analysis of more than 5000 protein structures (*30*). Lastly, considering both homo- and heteromultimers, almost 90% of soluble proteins were found to be part of at least one complex, a figure similar to values estimated in yeast (*3*, *4*). This further consolidates the view that exhaustive organization into complexes is a general property of proteomes in bacteria and eukaryotes.

Characteristics of *M. pneumoniae* protein complexes. Overall, more than half of the identified complexes were not previously described. We also found new components in previously known complexes: The data set contains 126 proteins with previously unknown or conflicting functional annotation. For example, complex membership identifies Mpn426, previously annotated as a P115 homolog, as the missing Smc (structural maintenance of chromosomes) DNA-binding subunit of the cohesin-like complex (complex 40, fig. S7B and table S4) (*28*). This complex also contains the adenosine triphosphate (ATP)–dependent protease Lon (Mpn332) that binds DNA and regulates chromosome replication (*31*). The observed physical association between Lon and Smc and the observation that Lon expression increases concomitant with Smc degradation at the onset of the stationary phase (fig. S7B) (*28*) suggest that Smc might be a target of this protease. The existence of a native complex including Lon, ScpA (Mpn300), and P115 is further supported by the observation that these three proteins co-elute during gel filtration chromatography (fig. S7B). We also identified known eukaryotic complexes such as those including several glycolytic enzymes (GEs) that have been discovered at eukaryotic plasma membrane, where they locally produce ATP (table S5). We observed similar assemblies in *M. pneumoniae* (complexes 12 and 45; table S4), which suggests that this function is conserved in bacteria.

Comparison of methods for estimation of proteome organization. We overlaid the protein complex data with complementary large-scale data sets that have been previously used to deduce physical interactions (Fig. 2A). Only 48% of the TAP interactions within complexes were found in any existing data set; 359 associations were only identified by TAP-MS (Fig. 2A). Even in the worst-case scenario, where we consider the upper limit of the estimated false-positives rate (20% = 100% to 80% accuracy) and assume that false positives are completely excluded from the other data sets, we estimate at least 220 previously unknown true associations were identified here. Overlap with interactions inferred from genome organization or gene expression was particularly low: Only 7% of the high-confidence interactions are between gene products from the same operon, and only 18% were consistently

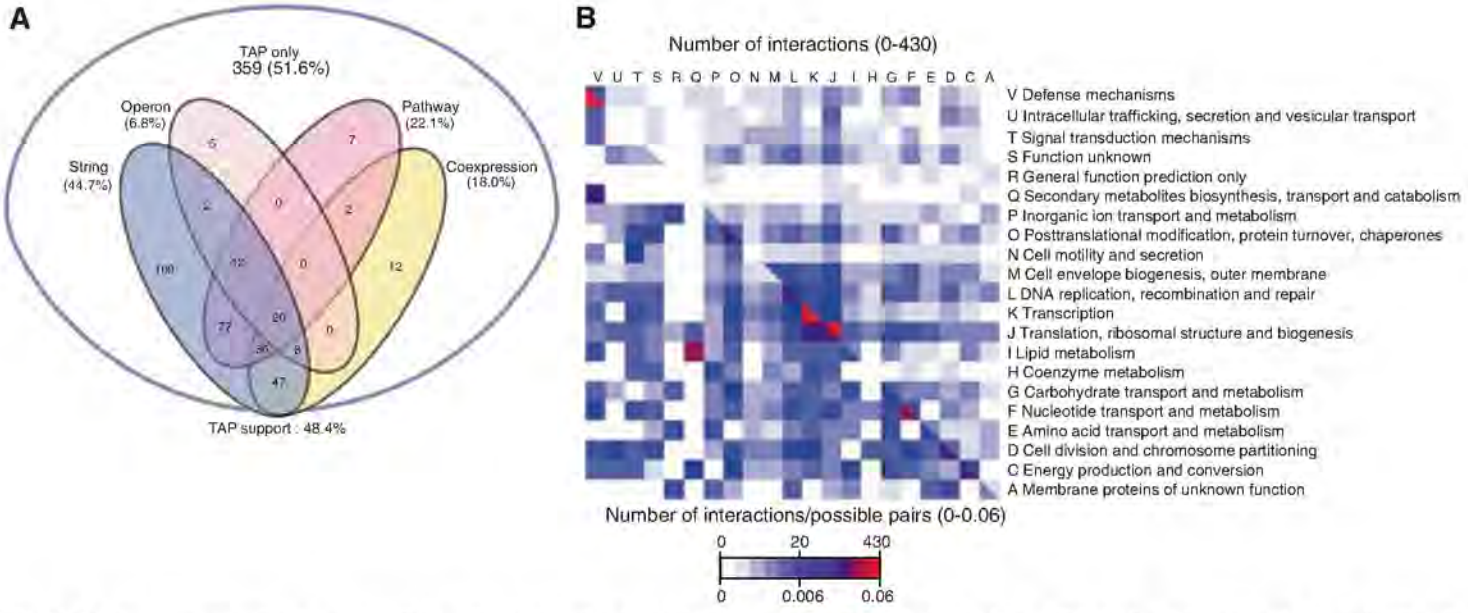


Fig. 2. Proteome organization is only partially reflected by other biological data sets. (**A**) General overlap between TAP and interactions inferred from other data sets: coexpression (*24*, *28*), operons (*24*), STRING (*21*), and pathways (*48*). Numbers refer to the interacting pairs within the different data sets. The fraction of TAP interactions that cluster into complexes and are covered by other data sets is given between brackets. For TAP-interacting protein pairs the cutoff was set at 80% accuracy. Cutoffs for other data sets were optimized for coverage (accuracies from 40 to 100%). (**B**) Frequent functional cross-talk in the protein complex data set. All proteins within high confidence pairs were functionally annotated according to the COG (Clusters of Orthologous Groups of Proteins) database (*49*). Boxed areas are colored proportionally to the number of interactions linking two functional classes. The scales represent the total (top) and normalized (bottom) number of interactions (*23*). Category Q (secondary metabolites) contains only two proteins. The category most frequently linked is J (translation) with itself; however, it contains the highest number of proteins. The highest proportion of interactions is between proteins within category K (transcription).

coexpressed (*24*). This implies that temporal or conditional regulation of complex formation is analogous to that for eukaryotes, in which different components are expressed at different times (*1*). For example, the four known subunits of the RNA polymerase are in three operons, and their transcription profiles correlate with two different gene expression groups along the growth curve (*24*, *28*). With current knowledge, only a small fraction of proteome organization can be inferred from analysis of the genomes or transcriptional data, making proteomics studies critical for understanding prokaryotic systems.

The *M. pneumoniae* protein complex network reveals substantial cross-talk. About a third of the heteromultimeric complexes in *M. pneumoniae* have extensive physical interconnections that suggest proteins participate in different cellular processes (Fig. 2B). These reflect protein multifunctionality (see below) and organization into at least 35 larger assemblies, sometimes hinting at physical, possibly temporal, associations of sequential steps in biological processes (table S4). For example, we reconstituted major parts of the ribosome from the interaction screen and saw extensive cross-talk with RNA polymerase (Fig. 3A). This higher-level association was unaffected by ribonuclease (RNAse) and deoxyribonuclease (DNAse) treatments, which suggests that protein-protein rather than protein–nucleic acid interactions were involved (fig. S5). The TAP-MS data were consistent with gel filtration results showing that the RNA polymerase α subunit, RpoA (Mpn191), and the ribosomal protein RpsD (Mpn446) co-elute with high apparent molecular sizes (Fig. 3A). These observations are further supported by the genome organization, where the rpoA gene is localized in and co-regulated with a ribosomal operon (*24*). This network provides a molecular model for the coupling of transcription and translation proposed in bacteria (*32*) and the direct involvement of ribosomal proteins in transcriptional regulation (*33*). The same assembly also includes translational initiation factors InfA (Mpn187), InfB (Mpn155), and InfC (Mpn115), which are part of the 30*S* initiation complex, as well as elongation factors Tuf (Mpn665) and Tsf (Mpn631), suggesting that we have captured sequential steps in a pathway running from transcription to translation.

Functional reuse and modularity of protein complexes. Genome-wide screens in eukaryotes show that proteins often participate in more than one complex, an attribute that has been proposed to account for protein multifunctionality, pleiotropy, and moonlighting (*34*). We defined a multifunctionality index that measures the tendency of proteins to associate with more than one complex (*16*). This index is based on frequency with which pairs of proteins were found associated in our set of purifications and is insensitive to the clustering parameters. We found 156 multifunctional proteins (table S7), covering 54% of *M. pneumoniae* proteins that are currently known to be multifunctional in the literature (table S8). We also compared our results with a set of multifunctional enzymes that catalyze different enzymatic reactions (*28*), and the overlap was smaller (32%). Our analysis captured distinct mechanisms for

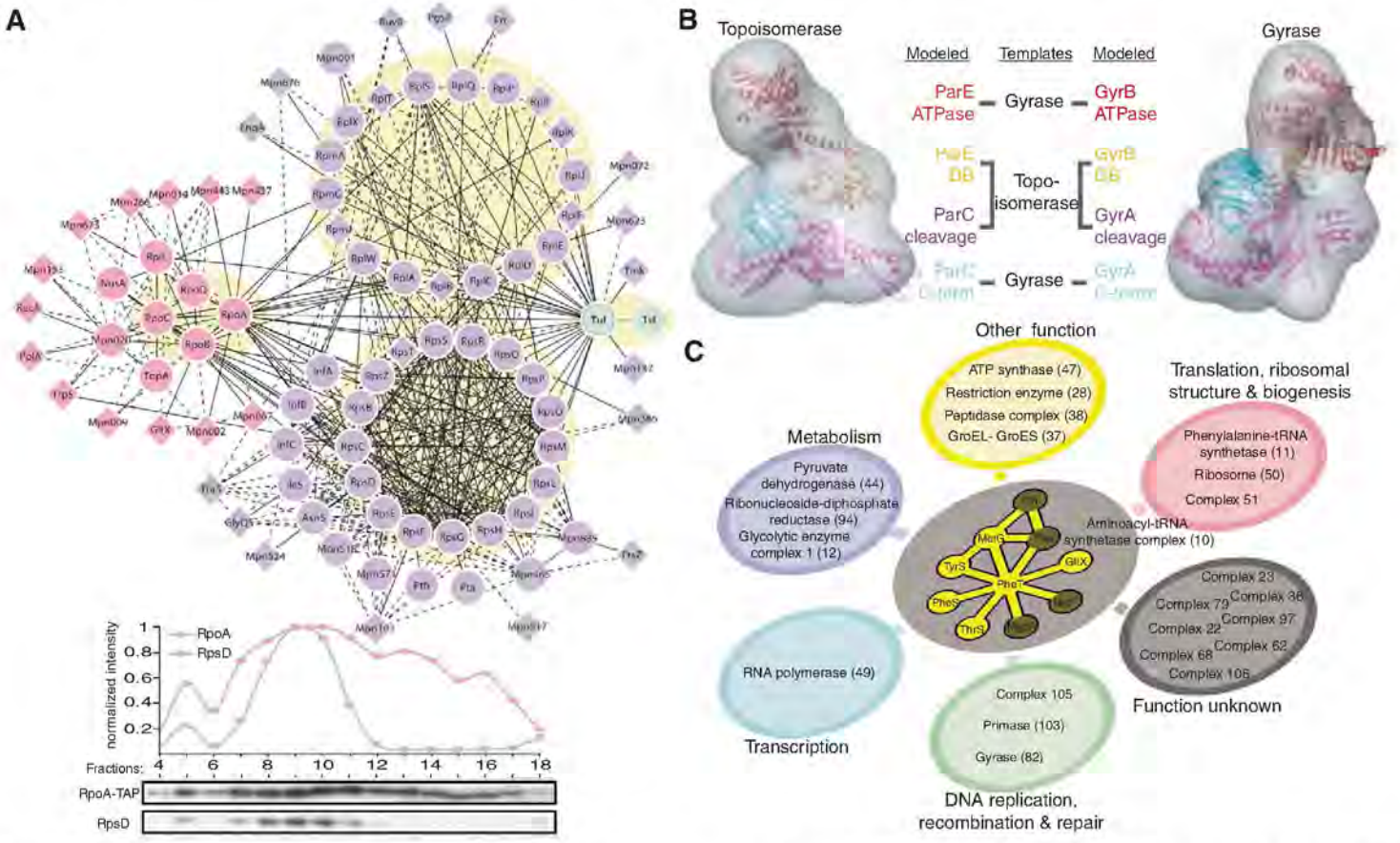


Fig. 3. Higher level of proteome organization. (**A**) The RNA polymerase–ribosome assembly. Core components are represented by circles, attachments by diamonds. The line attribute corresponds to socio-affinity indices: dashed lines, 0.5 to 0.86; plain lines, >0.86. Color code and shaded yellow circles around groups of proteins refer to individual complexes: RNA polymerase (pink), ribosome (purple), and translation elongation factor (green). The bottom graph shows that the ribosomal protein RpsD (23 kD) and the α subunit of the RNA polymerase, RpoA-TAP (57 kD), co-elute in high molecular weight fractions (MD range) during gel filtration chromatography. (**B**) DNA topoisomerase (diameter ~ 12 nm) is a heterodimer in bacteria: ParE (ATPase and DNA binding domains) and ParC (cleavage and C-terminal domains). The interaction between ParE-DNA–binding and ParC–cleavage domains was modeled by using yeast topoisomerase II as a template [Protein Data Bank (PDB) code 2rgr], and ParE-ATPase and ParC–C-terminal domains were modeled separately on structures of gyrase homologs (PDB 1kij and 1suu). All four domains were fitted into the electron microscopy density. Gyrase (~12 nm) is similarly split in bacteria into GyrA/GyrB, which are paralogs of ParE/ParC, and was modeled and fitted by using PDB 1bjt as a template for the GyrB-DNA–binding and GyrA-cleavage domains interaction. (**C**) Protein multifunctionality in *M. pneumoniae* illustrated with the AARS complexes.

multifunctionality that imply the combinatorial use of gene products in different contexts, for different functions.

For example, GyrA (Mpn004) is a component of the DNA gyrase complex that introduces negative supercoils into DNA, and ParE (Mpn122) is a member of the topoisomerase IV complex, which decatenates DNA (*35*). Besides well-documented interactions within their respective complexes (complexes 17 and 82, table S4), GyrA and ParE were also found to stably associate with each other (complex 102, table S4). Single-particle electron microscopy and comparative modeling (fig. S6) showed that DNA topoisomerase and DNA gyrase have related overall shapes, as expected from their functional similarity, and also support the notion that they might be able to interchange subunits (Fig. 3B). In eukaryotes, ParE and ParC (Mpn123) are fused into one single polypeptide. In bacteria, the possibility for the split ParE and ParC to contribute to different complexes might represent a parsimonious way of generating functional diversity and also robustness to mutations with a set of paralogous proteins.

Another example is a complex containing a cluster of five different aminoacyl transfer RNA (tRNA) synthetases (AARSs) (complex 10, Fig. 3C and table S4). In eukaryotes and archaea, AARSs form macromolecular complexes that improve aminoacylation efficiency by channeling substrates to ribosomes (*36*, *37*). These assemblies also act as reservoirs of AARSs that additionally exert a range of noncanonical regulatory functions in transcription, metabolism, and signaling (*38*). The existence in bacteria of big multi-AARS complexes is controversial; the most recent review advocates assembly in binary complexes that are functionally involved in tRNA metabolism and editing (*39*). Our results suggest that higher-order multi-AARS complexes might also exist in bacteria. We also found several AARSs in other complexes involved in functions as diverse as translation, transcription, DNA replication, and metabolism (Fig. 3C).

Structural anatomy of *M. pneumoniae*. Because of their small genome size, bacteria from the genus *Mycoplasma* have attracted attention as model organisms for structural genomics (*40*). We used these data to populate our protein complex network with structural information. Sequence

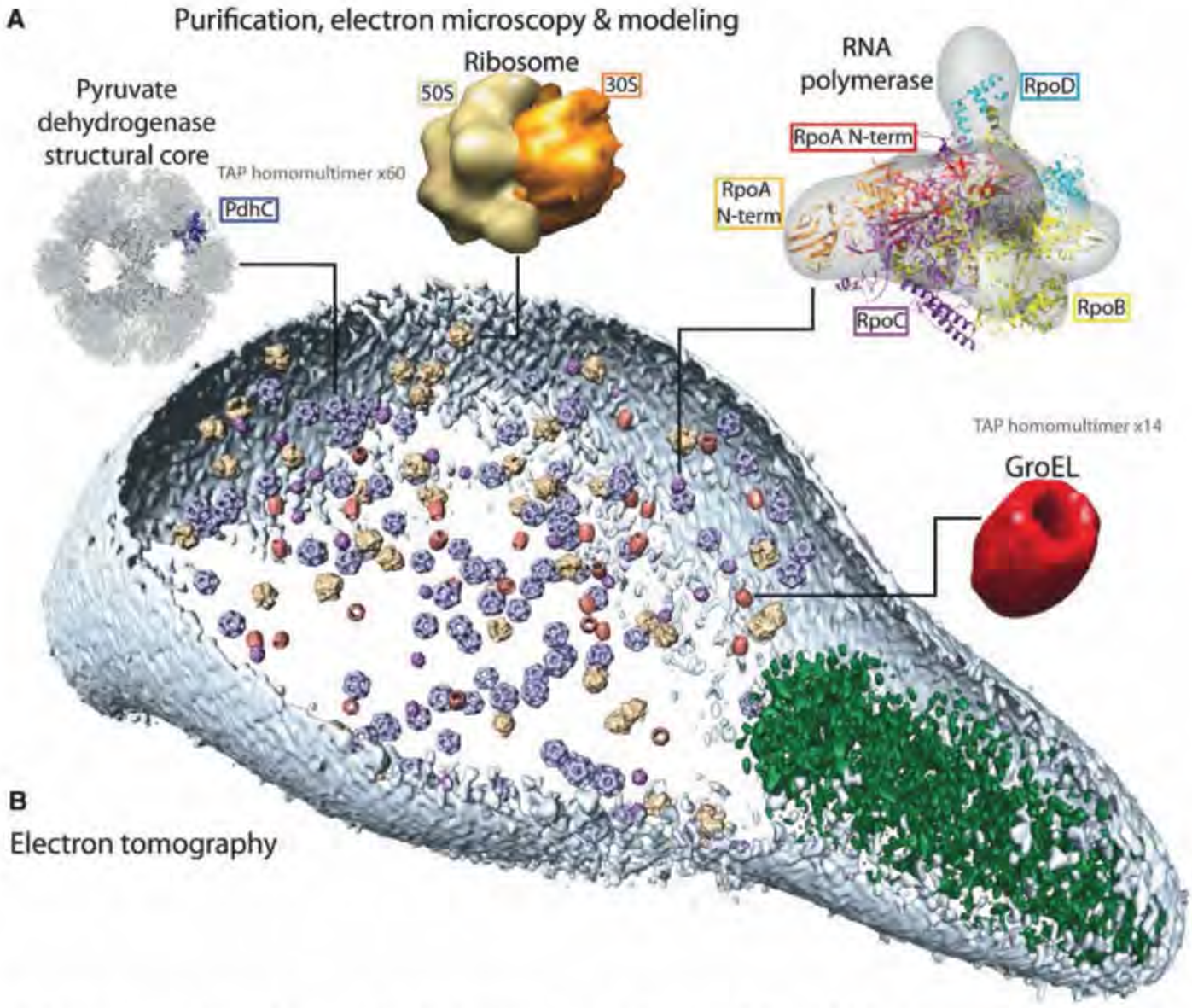


Fig. 4. From proteomics to the cell. By a combination of pattern recognition and classification algorithms, the following TAP-identified complexes from *M. pneumoniae*, matching to existing electron microscopy and x-ray and tomogram structures (**A**), were placed in a whole-cell tomogram (**B**): the structural core of pyruvate dehydrogenase in blue (~23 nm), the ribosome in yellow (~26 nm), RNA polymerase in purple (~17 nm), and GroEL homomultimer in red (~20 nm). Cell dimensions are ~300 nm by 700 nm. The cell membrane is shown in light blue. The rod, a prominent structure filling the space of the tip region, is depicted in green. Its major structural elements are HMW2 (Mpn310) in the core and HMW3 (Mpn452) in the periphery, stabilizing the rod (*42*).The individual complexes (A) are not to scale, but they are shown to scale within the bacterial cell (B).

similarity searches and comparative modeling provided structures for 484 *M. pneumoniae* proteins (70% of the genome) and 340 proteins in the network. There were also structural templates to construct models for 153 binary interactions (Fig. 1) covering 29 heteromultimeric and 57 homomultimeric complexes (table S6). These data can be used both to study particular interactions or complexes (Fig. 3B and fig. S7A) and to infer general correlations. Structural interfaces are particularly illuminating for the multifunctional proteins. When structural models are available for multiple interactions with a common protein, analysis of the interfaces can suggest whether the interactions are mutually exclusive (same binding sites) or compatible (different sites) (*41*). We observed that multifunctional proteins generally tend to accommodate more ligands per interacting interface ($P = 0.003$), consistent with the view that multifunctionality engages mutually exclusive interactions. For example, the protein P115 (Mpn426) has six distinct interfaces, each of which has several mutually exclusive interaction partners.

Having assembled a repertoire of structural information, the next logical step is to map these networks and protein complexes in their native environment, the cell. For this purpose, we performed cryogenic electron tomography of 26 entire *M. pneumoniae* cells (*42*) (fig. S8). We used pattern recognition techniques to generate probability maps for complexes selected from the larger ones in *M. pneumoniae* (Fig. 4) because larger complexes are more likely to be identified. After a thorough classification considering missing data, low signal-to-noise ratio, and known spatial proximities of different subcomplexes, we generated maps for the ribosome, the chaperone GroEL (Mpn573), the structural core of the pyruvate dehydrogenase (PdhC, Mpn391, homomultimer), and RNA polymerase, with a minimal number of false positives (Fig. 4). These large complexes are excluded from the tip, an organelle required for the attachment to epithelial cells, illustrating that even in a simple, minimal bacteria the proteome is spatially organized (*42*). Within the cell bodies, we could not find substantial proximities or patterns among the different complexes. In contrast to *E. coli* that contains a compact nucleoid forming an exclusion area in the cell center (*43*), circular DNA in *M. pneumoniae* is apparently uniformly distributed (*44*). We estimated the average number of complexes per cell to be 140 for the ribosome, 100 for GroELs, 100 for pyruvate dehydrogenase, and 300 for RNA polymerase. For the ribosome and GroEL, we also quantified complex abundances by Western blotting (fig. S9). For both, the numbers derived from Western blot were in the range of those estimated from the tomograms. This adds to the emerging view that the mapping of macromolecular structures into entire-cell tomograms (*45*), even though still challenging, is a powerful strategy when combined with unbiased large-scale complex purification. It opens the way to more general charting of cellular networks in entire-cell tomograms.

Conclusions. Our genome-scale screen for soluble complexes in a bacterium provides a valuable resource for the functional annotation of many genes whose biological roles in prokaryotic or parasitic cells are elusive. The coverage of known complexes leads to an estimate of some 200 molecular machines in *M. pneumoniae*. The study allows estimation of unanticipated proteome complexity for an apparently minimal organism that could not be directly inferred from its genome composition and organization or from extensive transcriptional analysis. Organisms with small genomes are the most tractable for systems biology, and the biochemical data set, proteome-wide spectra, ORFome, and collection of TAP-expressing *M. pneumoniae* strains will provide an extremely useful resource for this community. Comparison to both more complex bacteria and to even smaller ones, such as *M. genitalium* with 485 annotated protein-coding genes (*46*), should reveal additional systemic features associated with genome streamlining.

With protein structures available for about three-quarters of its ORFs, either directly from structural genomics efforts (*40*) or indirectly inferred by homology, *M. pneumoniae* has been extensively studied. We demonstrated that we can integrate data sets of biochemically determined complexes with structural information to approximate the three-dimensional organization of proteins into functional molecular machines. These models can then be mapped in entire cell tomograms, providing a three-dimensional view of cellular proteomes and interactomes (*47*); ultimately whole-cell models will benefit studies of biological function and disease.

References and Notes

1. U. de Lichtenberg, L. J. Jensen, S. Brunak, P. Bork, *Science* **307**, 724 (2005).
2. G. Rigaut *et al.*, *Nat. Biotechnol.* **17**, 1030 (1999).
3. A. C. Gavin *et al.*, *Nature* **440**, 631 (2006).
4. N. J. Krogan *et al.*, *Nature* **440**, 637 (2006).
5. K. Tarassov *et al.*, *Science* **320**, 1465 (2008); published online 7 May 2008 (10.1126/science.1153878).
6. J. C. Rain *et al.*, *Nature* **409**, 211 (2001).
7. J. R. Parrish *et al.*, *Genome Biol.* **8**, R130 (2007).
8. L. Terradot *et al.*, *Mol. Cell. Proteomics* **3**, 809 (2004).
9. G. Butland *et al.*, *Nature* **433**, 531 (2005).
10. M. Arifuzzaman *et al.*, *Genome Res.* **16**, 686 (2006).
11. P. Hu *et al.*, *PLoS Biol.* **7**, e96 (2009).
12. K. B. Waites, D. F. Talkington, *Clin. Microbiol. Rev.* **17**, 697 (2004).
13. R. Himmelreich *et al.*, *Nucleic Acids Res.* **24**, 4420 (1996).
14. T. Dandekar *et al.*, *Nucleic Acids Res.* **28**, 3278 (2000).
15. J. I. Glass *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 425 (2006).
16. Materials and methods are available as supporting material on *Science* Online.
17. D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell, *Electrophoresis* **20**, 3551 (1999).
18. E. Gasteiger *et al.*, in *The Proteomics Protocols Handbook*, J. M. Walker, Ed. (Humana, Totowa, NJ, 2005), pp. 571–607.
19. K. A. Resing *et al.*, *Anal. Chem.* **76**, 3556 (2004).
20. F. Desiere *et al.*, *Nucleic Acids Res.* **34**, D655 (2006).
21. L. J. Jensen *et al.*, *Nucleic Acids Res.* **37**, D412 (2009).
22. M. E. Sowa, E. J. Bennett, S. P. Gygi, J. W. Harper, *Cell* **138**, 389 (2009).
23. C. von Mering *et al.*, *Nature* **417**, 399 (2002).
24. M. Güell *et al.*, *Science* **326**, 1268 (2009).
25. A. C. Gavin *et al.*, *Nature* **415**, 141 (2002).
26. H. Berman, K. Henrick, H. Nakamura, *Nat. Struct. Biol.* **10**, 980 (2003).
27. P. Zuber, *J. Bacteriol.* **186**, 1911 (2004).
28. E. Yus *et al.*, *Science* **326**, 1263 (2009).
29. K. S. Murakami, S. Masuda, S. A. Darst, *Science* **296**, 1280 (2002).
30. E. D. Levy, E. Boeri Erba, C. V. Robinson, S. A. Teichmann, *Nature* **453**, 1262 (2008).
31. R. Wright, C. Stephens, G. Zweiger, L. Shapiro, M. R. Alley, *Genes Dev.* **10**, 1532 (1996).
32. J. Gowrishankar, R. Harinarayanan, *Mol. Microbiol.* **54**, 598 (2004).
33. M. Torres, C. Condon, J. M. Balada, C. Squires, C. L. Squires, *EMBO J.* **20**, 3811 (2001).
34. J. Hodgkin, *Int. J. Dev. Biol.* **42**, 501 (1998).
35. E. L. Zechiedrich, N. R. Cozzarelli, *Genes Dev.* **9**, 2859 (1995).
36. M. Praetorius-Ibba, C. D. Hausmann, M. Paras, T. E. Rogers, M. Ibba, *J. Biol. Chem.* **282**, 3680 (2007).
37. S. V. Kyriacou, M. P. Deutscher, *Mol. Cell* **29**, 419 (2008).
38. S. G. Park, P. Schimmel, S. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11043 (2008).
39. C. D. Hausmann, M. Ibba, *FEMS Microbiol. Rev.* **32**, 705 (2008).
40. S. H. Kim *et al.*, *J. Struct. Funct. Genomics* **6**, 63 (2005).
41. P. M. Kim, L. J. Lu, Y. Xia, M. B. Gerstein, *Science* **314**, 1938 (2006).
42. A. Seybert, R. Herrmann, A. S. Frangakis, *J. Struct. Biol.* **156**, 342 (2006).
43. M. Thanbichler, L. Shapiro, *Nat. Rev. Microbiol.* **6**, 28 (2008).
44. S. Seto, G. Layh-Schmitt, T. Kenri, M. Miyata, *J. Bacteriol.* **183**, 1621 (2001).
45. A. Al-Amoudi, D. C. Diez, M. J. Betts, A. S. Frangakis, *Nature* **450**, 832 (2007).
46. D. G. Gibson *et al.*, *Science* **319**, 1215 (2008); published online 23 January 2008 (10.1126/science.1151721).
47. P. Bork, L. Serrano, *Cell* **121**, 507 (2005).
48. M. Kanehisa *et al.*, *Nucleic Acids Res.* **36**, D480 (2008).
49. R. L. Tatusov, E. V. Koonin, D. J. Lipman, *Science* **278**, 631 (1997).

50. We are grateful to M. Wilm, T. Franz, F. Thommen, E. Dalton, M. Schulz, E. Sawa, M. Diepholz, E. Pirkl, A. Seybert, C. Davis, J. Stülke, Gavin's and Bork's groups, and the EMBL Proteomic and Gene Core Facilities for expert help and discussion. This work is in part supported by the European Commission 6th and 7th Framework Integrated Projects "3D-Repertoire" and "Prospects," respectively; SystemsX.ch, the Swiss initiative for systems biology; the Netherlands Organization for Scientific Research (NWO); the Foundation Marcelino Botín; the Spanish Ministry of Education and Science (MEC)–Consolider; and the European Research Council. The data set has been submitted to the International Molecular Exchange Consortium (http://imex.sf.net) through IntAct (pmid is 17145710; identifier is IM-11644). The electron microscopy maps have been submitted to the Electron Microscopy Data Bank (www.ebi.ac.uk/pdbe-srv/emsearch/) (identification codes EMD-1637, EMD-1638, and EMD-1639).

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1235/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S8

15 May 2009; accepted 2 October 2009
10.1126/science.1176343

Directed Transport of Atoms in a Hamiltonian Quantum Ratchet

Tobias Salger,[1]* Sebastian Kling,[1] Tim Hecking,[1] Carsten Geckeler,[1] Luis Morales-Molina,[2] Martin Weitz[1]

Classical ratchet potentials, which alternate a driving potential with periodic random dissipative motion, can account for the operation of biological motors. We demonstrate the operation of a quantum ratchet, which differs from classical ratchets in that dissipative processes are absent within the observation time of the system (Hamiltonian regime). An atomic rubidium Bose-Einstein condensate is exposed to a sawtooth-like optical lattice potential, whose amplitude is periodically modulated in time. The ratchet transport arises from broken spatiotemporal symmetries of the driven potential, resulting in a desymmetrization of transporting eigenstates (Floquet states). The full quantum character of the ratchet transport was demonstrated by the measured atomic current oscillating around a nonzero stationary value at longer observation times, resonances occurring at positions determined by the photon recoil, and dependence of the transport current on the initial phase of the driving potential.

Transport phenomena are at the heart of many problems in physics, chemistry, and biology. An intriguing example is the ratchet effect, where in the absence of a biased force, a directed current of particles is produced along a periodic structure (*1*). Ratchets underlie a model for the operational principle of biological motors, tiny biological engines that transform the energy produced in chemical reactions into unidirectional mechanical motion (*2*). More generally, ratchets also provide models for microscopic machinery and allow for a directed current in the presence of thermal noise far from thermal equilibrium (*3*). Experimentally, ratchet transport has been studied in solid-state systems where some quantum effects were observed (*4*, *5*) and, more recently, in atomic physics systems (*6*–*9*).

Cold atoms in optical lattices can be used to study quantum transport in the effective absence of dissipation when the optical fields are detuned far from resonance. In notable experiments, tunneling oscillations and quantum resonances have been observed with moving lattice potentials (*10*–*12*). However, directed transport of particles that are initially at rest with respect to the lattice requires that the driving potential breaks certain symmetries (*13*). In one experiment, directed motion was achieved with delta-kicked atoms and attributed to asymmetric dynamical localization (*9*). In other experiments with laser-cooled atoms subject to a continuous drive, the relation between the appearance of transport and symmetry breaking has been investigated, and the scaling of transport versus dissipation was studied (*6*–*8*). Theoretical work has pointed out that ratchet transport should also be possible in Hamiltonian quantum systems, where no dissipative processes aid to break the time-inversion symmetry and, thus, ensure unidirectional atom transport (*14*–*16*). The interest in such proposed devices is that they would provide a blueprint for new quantum machinery (*17*).

We here report on evidence for atom transport in a Hamiltonian (nondissipative) quantum ratchet. A rubidium Bose-Einstein condensate (BEC) is exposed to a sawtooth-like optical lattice potential that is periodically modulated in time. When the driving lattice potential breaks spatial and temporal symmetries, directed transport of atoms is observed. For a flashing sawtooth potential (Fig. 1A), ratchet transport is possible for a classical particle undergoing Brownian motion. Directed transport proceeds by particles rolling down a potential hill, accumulating in a potential minimum and during the subsequent off-phase of the potential; thermal motion allows for transport from the location of one minimum to the next (*1*). In contrast to ratchets based on dissipative forces and thermal motion, relaxation processes cannot be used to ensure directed motion in a Hamiltonian quantum system. For a quantum system periodically modulated in time, the dynamics can be described in terms of eigenfunctions of the Hamiltonian: $H(t) = p^2/2m + V(x, t)$ with $V(x, t) = V(x, t + T)$, where p represents atomic momentum, m is atomic mass, V is driving potential, x is position, t is time, and T denotes the temporal periodicity. These so-called Floquet states are the analogs of energy eigenstates in time-periodic quantum systems (*18*). Each of these states carries a distinct average momentum; a representation of two Floquet states in our flashing ratchet system is shown in Fig. 1B. We illustrate the Floquet states of nonzero average momentum as transporting conveyor belts that move in different directions along the lattice axis (Fig. 1C). After loading of an atomic BEC into the lattice, the atomic wave function can be expressed as a coherent superposition of different Floquet eigenstates, where the relative phases and amplitudes are determined by the overlap to the BEC wave function. Directly after the loading, the relative phases are such that

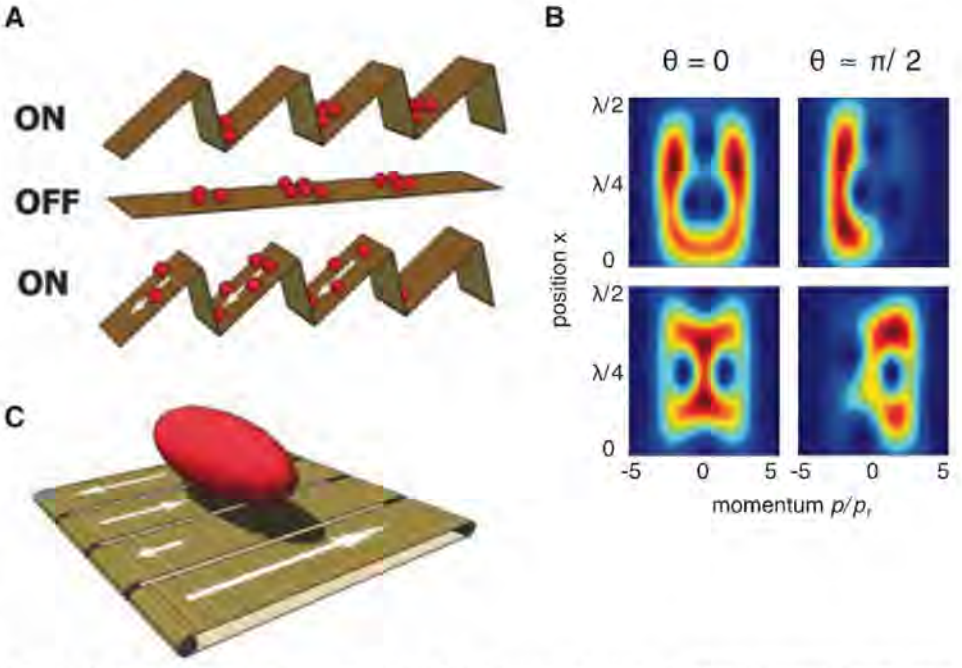


Fig. 1. (**A**) Operation of a Brownian ratchet with classical particles. (**B**) Quasi-classical Husimi representation of the phase space for two different Floquet eigenstates (top and bottom) of a Hamiltonian quantum ratchet $V(x, t) = [V_1\cos(2kx) + V_2\cos(4kx + \phi)] \cdot [A_1\sin^2(\omega t/2) + A_2\sin^2(\omega t + \theta/2)]$. The parameters were as follows: $\omega/8\omega_r = 1.05$, $A_1 = 0.78$, $A_2 = 0.22$, $V_1 = 5.3E_r$, $V_2 = 1.3E_r$, and $\phi \cong \pi/2$, where $E_r = \hbar\omega_r$ is the recoil energy, for $\theta \cong 0$ (left) and $\theta \cong \pi/2$ (right). In the latter case, a desymmetrization of Floquet states is achieved. (**C**) Graphical scheme illustrating the loading of an atomic BEC onto Floquet eigenstates of the flashing ratchet potential, visualized here as conveyor belts to represent the associated atom transport.

[1]Institut für Angewandte Physik, Wegelerstrasse 8, 53115 Bonn, Germany. [2]Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago, Chile.

*To whom correspondence should be addressed. E-mail: salger@iap.uni-bonn.de

the total momentum of the quantum state equals that of the BEC (that is, the atoms are still at rest).

Subsequently, phase factors determining the wave function temporal evolution become important, and we expect oscillatory motion to arise from the beating between eigenstates. The corresponding eigenfrequencies are incommensurate, and, for larger times, the atomic momentum oscillates around an average value determined by the weighted superposition of the momenta of the populated eigenstates. Thus, an asymptotic directed transport can only be obtained when the populated eigenstates have a nonzero average momentum, which requires a breaking of spatiotemporal symmetries. Besides the Floquet oscillation around a nonzero average value, two further benchmarks of Hamiltonian quantum ratchets are a dependence on the initial time of the external driving and the precise vanishing of the quantum-ratchet transport at symmetry-restoring points (*16*). None of these salient features have been experimentally validated in previous works studying atomic quantum-ratchet transport.

In our experiment, an atomic rubidium BEC is loaded into a flashing ratchet potential of the form $V(x, t) = V(x) \cdot A(t)$, with $V(x) = V_1\cos(2kx) + V_2\cos(4kx + \phi)$, where A comprises the amplitude dependence of the driving lattice potential V_1 and V_2 denote the amplitudes of the spatial lattice harmonics with periodicities $\lambda/2$ and $\lambda/4$, respectively, which create the ratchet potential; k is the wave vector of the driving optical field and equals $2\pi/\lambda$, where λ is the wavelength of the driving laser; and ϕ is the relative phase between the lattice harmonics. The fundamental spatial frequency of periodicity $\lambda/2$ is generated by a usual standing wave formed with two counterpropagating fields; in our experiment, λ is 1.5 nm (10^5 natural linewidths) detuned to the red of the D2 line. A lattice with spatial periodicity $\lambda/4$ is realized with the dispersion of Doppler-sensitive, four-photon Raman transitions between Zeeman ground-state sublevels (*19*–*21*), driven by the same laser. The temporal modulation of the potential is described with a biharmonic function $A(t) = A_1\sin^2[\omega(t - t_0)/2] + A_2\sin^2[\omega(t - t_0) + \theta/2]$, where A_1 and A_2 denote the amplitudes of the harmonics, respectively, $\omega = 2\pi/T$ is the modulation frequency, θ is the relative phase between harmonics, and t_0 is the initial time of the modulation, with $A(t < t_0) = 0$. As shown in (*13*), a net transport in the flashing ratchet system requires one to break certain temporal and spatial symmetries [see the supporting online material (*22*)]. These conditions are fulfilled if $A_1, A_2 \neq 0$ and $\theta \neq 0, \pm\pi$ along with $V_1, V_2 \neq 0$ and $\phi \neq 0, \pm\pi$. After the interaction with the flashing ratchet potential, the atoms freely expand for a 15-ms period, after which an absorption image is recorded. This time-of-flight (TOF) technique can be used to analyze the atomic-velocity distribution.

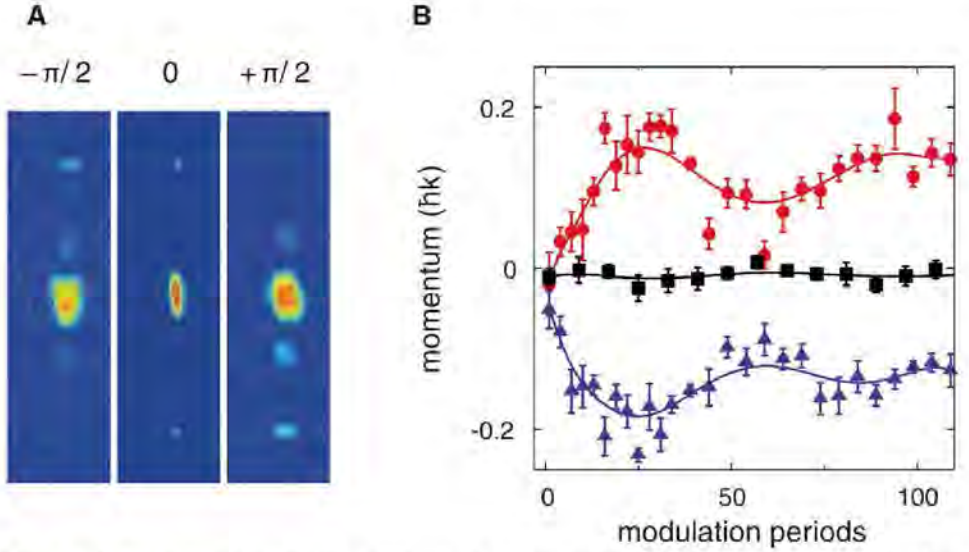


Fig. 2. (**A**) TOF images showing the atomic-velocity distribution after 26 modulation periods of the ratchet potential for a relative phase between temporal harmonics of $\theta = -\pi/2$ (left), $\theta = 0$ (middle), and $\theta = \pi/2$ (right). The modulation period was $T = 2\pi/\omega \cong 32.1$ μs, for which $\omega \cong 8\omega_r$. Other parameters of the driving potential were $V_1 = 6E_r$, $V_2 = 1.6E_r$, a phase $\phi = \pi/2$, and amplitudes of temporal harmonics $A_1 = 0.75$ and $A_2 = 0.40$. These values were chosen to maximize the directed transport. (**B**) Mean atomic momentum as a function of the number of modulation periods for $\theta = -\pi/2$ (blue triangles), $\theta = 0$ (black squares), and $\theta = \pi/2$ (red dots). The data points are fitted with a spline function. Error bars indicate SE of the mean.

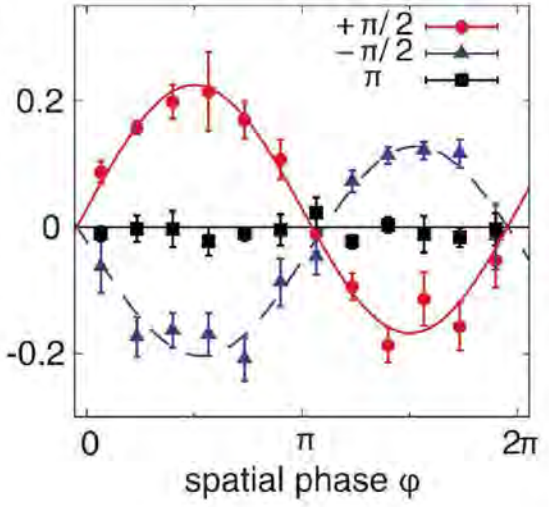


Fig. 3. Mean atomic momentum versus the phase ϕ between spatial lattice harmonics for three different values of the phase between temporal harmonics of the flashing ratchet: (i) $\theta = -\pi/2$ (blue triangles), (ii) $\theta = 0$ (black squares), and (iii) $\theta = \pi/2$ (red dots). To guide the eye, the data points have been fitted with sinusoidal curves. Parameters are the same as in Fig. 2A.

For a more detailed analysis of the transport, the mean atomic momentum along the lattice axis is calculated from such TOF images. Corresponding results are shown in Fig. 2B as a function of the interaction time with the flashing ratchet potential. The BEC, prepared initially at rest with respect to the lattice reference frame, was accelerated to positive (or negative) momentum values for a phase between the two driving harmonics of $\theta = \pi/2$ (or $\theta = -\pi/2$, respectively). The atoms continued to accelerate until a velocity of ~$0.15\hbar k/m$ (–$0.15\hbar k/m$ for opposite temporal phase) was reached after 30 periods of driving. For larger interaction times, an oscillation around a nonzero mean value was observed. The oscillation is attributed to the interference between discrete Floquet eigenstates, which is in agreement with theoretical predictions for Hamiltonian quantum-ratchet transport (*16*).

Results for the dependence of the atom transport on the spatial symmetry of the ratchet are displayed in Fig. 3, showing the mean momentum as a function of the spatial phase between lattice harmonics for different values of the temporal

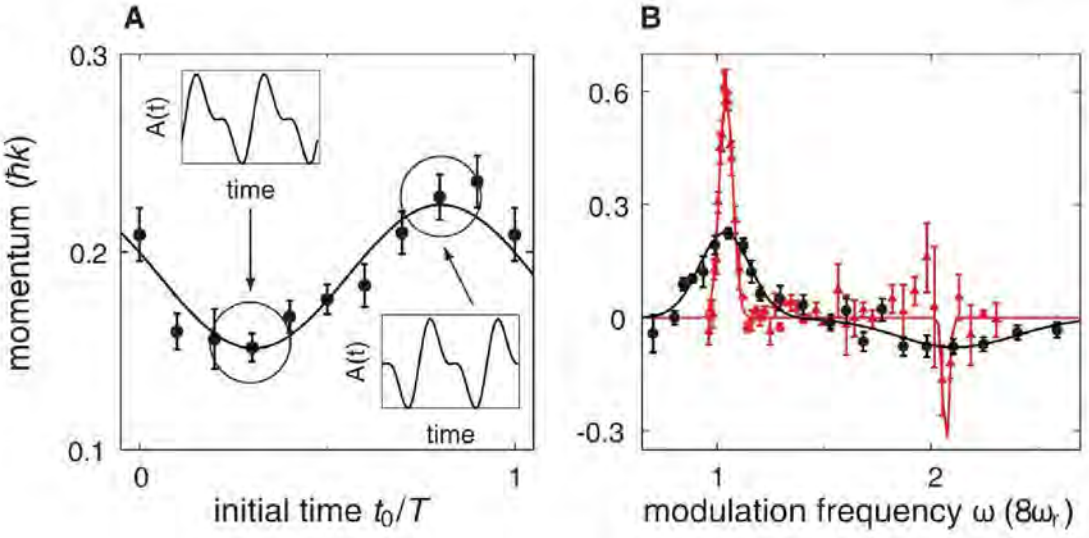


Fig. 4. **(A)** Mean atomic momentum as a function of the initial time of modulation t_0 of the flashing ratchet. **(B)** Mean atomic momentum as a function of the modulation frequency ω (black dots). The red triangles give data recorded with additional Bragg selection of a narrowed atomic-velocity slice. The data points have been fitted with two Gaussian functions to guide the eye. Error bars indicate SE of the mean.

phase. As expected, the directed transport is maximized when both phases ϕ and θ are near $\pm\pi/2$. The breaking of spatiotemporal symmetries is then largest, and the Floquet eigenstates are desymmetrized. In contrast, when $\phi = 0$ or $\theta = 0$, no clear directed transport is observed, which shows that no significant drifts were present in our experimental apparatus.

An important prediction for Hamiltonian quantum ratchets is a dependence of the current on the initial time t_0 (*16*). A variation in t_0 corresponds to a phase shift within the driving cycle. Figure 4A shows experimental data for the atom transport as a function of this initial time, and the insets indicate the form of the modulation for two values of t_0. For a proper choice of the initial time, an enhancement of the directed transport by some 30% could be achieved. Because of ergodicity in the chaotic layer of phase space, no dependence of the transport on the initial time is expected in the classical Hamiltonian system or any dissipative ratchet system (*16*). In contrast, Hamiltonian quantum systems memorize the initial condition.

For a further test of the quantum nature of the effect, the dependence of the ratchet transport on the modulation frequency ω of the flashing ratchet was studied. A tuning of the modulation frequency allows for critical control of the position of the (Floquet) eigenstates of the system, and maximum transport is expected for strong mixing between such states, corresponding to a quantum resonance (*1*). Figure 4B shows measured data for the directed ratchet transport versus the drive frequency (see the black dots first). A strong principle resonance occurs near $\omega = 8\omega_r$, with $\omega_r = \hbar k^2/2m$ as the photon recoil frequency (at which the previously shown data sets were recorded), and a second resonance with oppositely directed transport occurs at twice this value. In a simplified atom-optics picture, the resonances can be understood by interpreting the flashing ratchet potential as a sequence of (temporally overlapping) optical beamsplitter pulses each diffracting atom waves, realizing a nested multiple-path atom interferometer. The atoms interfere in several families of wave packets, and a required condition for the different families to give a joint interference signal is that the path difference between adjacent paths reaches an integer multiple of 2π, from which the same resonance condition can be derived because of the presence of recoil phases quadratic in atomic momentum (*23*).

To test whether the width of the atomic-velocity distribution affects the linewidth of the observed resonances, we have, in a subsequent measurement, recorded data with reduced width of the initial atomic-velocity distribution. The atomic Doppler width for a two-photon transition after release of the interacting condensate atoms from the trap equals 10 kHz, which can be reduced to 3 kHz by Bragg selection of atoms, followed by a Raman transfer pulse. Transport data recorded with the velocity-selected atom source are shown by the red triangles in Fig. 4B versus drive frequency, which yields resonances with a narrower spectral width. The velocity selection also increases the magnitude of the mean momentum of the directed transport to a value of $0.6\hbar k/m$, which is close to the expected results (*16*). We attribute the increased transport to the velocity width of the Bragg-selected atom source fitting better into the narrow velocity-acceptance profile of the (Doppler-sensitive) quantum ratchet than that of the condensate atoms released directly from the trap.

The evidence for atom transport in a Hamiltonian quantum ratchet is obtained from (i) a saturation of the directed transport along with oscillations around the asymptotic current for an ensemble initially at rest, (ii) a dependence of the transport on the initial time of modulation, and (iii) discrete resonances in the modulation frequency near positions determined by the photon recoil, as predicted in (*16*). The results may be interpreted as a quantum motor that provides directed transport of atoms in the absence of a net directed force, with the use of operational principles that are fully quantum in nature. The results pave the way to a quantum motor delivering mechanical energy as a further step toward quantum machinery (*17*).

References and Notes

1. P. Hänggi, F. Marchesoni, *Rev. Mod. Phys.* **81**, 387 (2009).
2. F. Jülicher, A. Ajdari, J. Prost, *Rev. Mod. Phys.* **69**, 1269 (1997).
3. R. D. Astumian, P. Hänggi, *Phys. Today* **55**, 33 (2002).
4. H. Linke *et al.*, *Science* **286**, 2314 (1999).
5. J. B. Majer, P. Peguiron, M. Grifoni, M. Tusveld, J. E. Mooij, *Phys. Rev. Lett.* **90**, 056802 (2003).
6. M. Schiavoni, L. Sanchez-Palencia, F. Renzoni, G. Grynberg, *Phys. Rev. Lett.* **90**, 094101 (2003).
7. R. Gommers, S. Bergamini, F. Renzoni, *Phys. Rev. Lett.* **95**, 073003 (2005).
8. R. Gommers, S. Denisov, F. Renzoni, *Phys. Rev. Lett.* **96**, 240604 (2006).
9. P. H. Jones, M. Goonasekera, D. R. Meacher, T. Jeckheree, T. S. Monteiro, *Phys. Rev. Lett.* **98**, 073002 (2007).
10. D. A. Steck, W. H. Oskay, M. G. Raizen, *Science* **293**, 274 (2001); published online 5 July 2001 (10.1126/science.1061569).
11. W. K. Hensinger *et al.*, *Nature* **412**, 52 (2001).
12. I. Dana, V. Ramareddy, I. Talukdar, G. S. Summy, *Phys. Rev. Lett.* **100**, 024103 (2008).
13. S. Flach, O. Yevtushenko, Y. Zolotaryuk, *Phys. Rev. Lett.* **84**, 2358 (2000).
14. P. Reimann, M. Grifoni, P. Hänggi, *Phys. Rev. Lett.* **79**, 10 (1997).
15. H. Schanz, T. Dittrich, R. Ketzmerick, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **71**, 026228 (2005).
16. S. Denisov, L. Morales-Molina, S. Flach, P. Hänggi, *Phys. Rev. A* **75**, 063424 (2007).
17. A. V. Ponomarev, S. Denisov, P. Hänggi, *Phys. Rev. Lett.* **102**, 230601 (2009).
18. L. E. Reichl, *The Transition to Chaos in Conservative Classical Systems: Quantum Manifestations* (Springer, New York, 1992).
19. G. Ritt, C. Geckeler, T. Salger, G. Cennini, M. Weitz, *Phys. Rev. A* **74**, 063622 (2006).
20. T. Salger, C. Geckeler, S. Kling, M. Weitz, *Phys. Rev. Lett.* **99**, 190405 (2007).
21. P. R. Berman, B. Dubetsky, J. L. Cohen, *Phys. Rev. A* **58**, 4801 (1998).
22. See the supporting material on *Science* Online.
23. M. Weitz, T. Heupel, T. W. Hänsch, *Europhys. Lett.* **37**, 517 (1997).

24. We thank the Deutsche Forschungsgemeinschaft for financial support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1241/DC1
Materials and Methods
Fig. S1
References

23 July 2009; accepted 8 September 2009
10.1126/science.1179546

Visualizing the 3D Internal Structure of Calcite Single Crystals Grown in Agarose Hydrogels

Hanying Li,[1]* Huolin L. Xin,[2]* David A. Muller,[3] Lara A. Estroff[1]†

Single crystals are usually faceted solids with homogeneous chemical compositions. Biogenic and synthetic calcite single crystals, however, have been found to incorporate macromolecules, spurring investigations of how large molecules are distributed within the crystals without substantially disrupting the crystalline lattice. Here, electron tomography reveals how random, three-dimensional networks of agarose nanofibers are incorporated into single crystals of synthetic calcite by allowing both high- and low-energy fiber/crystal interface facets to satisfy network curvatures. These results suggest that physical entrapment of polymer aggregates is a viable mechanism by which macromolecules can become incorporated inside inorganic single crystals. As such, this work has implications for understanding the structure and formation of biominerals as well as toward the development of new high–surface area, single-crystal composite materials.

Biominerals, such as sea urchin skeletal parts and the calcite prisms in mollusk shells, are examples of organic/inorganic single-crystal composites with improved mechanical properties as compared with pure crystals (*1*–*7*). Sea urchin spines and plates are high-magnesium calcite single crystals, which incorporate not only micrometer-sized cellular tissue networks but also smaller intracrystalline proteins (*1*, *4*). Similarly, prisms from the calcitic layer of mollusk shells (such as *Atrina rigida*) are single crystals of calcite with an incorporated organic matrix composed of proteins and other biomacromolecules (*6*, *7*). Several techniques have been used to investigate how these macromolecules are distributed inside the single crystals without substantially disrupting the crystalline lattice. On the basis of coherence-length measurements on biogenic calcite crystals, Berman *et al.* suggested that intracrystalline proteins were concentrated at the mosaic block boundaries (*8*), whereas transmission electron microscopy (TEM) observation of sea urchin teeth has indicated that organic material may be isolated in nanometer-sized cavities (*5*). Synchrotron powder diffraction studies of a variety of biogenic calcite crystals indicate that the intra-crystalline organics induce anisotropic lattice distortions, suggesting that individual proteins may be incorporated as impurities in the crystalline lattice (*9*). In addition to biogenic crystals, synthetic calcite crystals also incorporate polymers when grown in the presence of polymer networks, porous membranes, or colloidal particles (*10*–*15*). Although it is evident that polymers are occluded in all of these crystals, the mode of distribution of the macromolecules within the crystals (such as individual molecules versus aggregates) and the structure of the polymer/crystal interface are still both poorly characterized and understood.

There are several challenges to the study of the internal structure of inorganic crystals with incorporated organic macromolecules at atomic resolution and in three dimensions by use of traditional EM and x-ray diffraction techniques. These difficulties arise from the small sizes of the incorporated organic molecules and aggregates, the radiation sensitivity (for example, electron damage) of the materials, and the lack of well-developed methods for preparing electron-transparent samples from micron-sized specimens suitable for high-resolution imaging. Aided by advances in focused ion beam (FIB) sample preparation technique, annular dark-field scanning TEM (ADF-STEM), and ADF electron tomography for three-dimensional (3D) imaging, we have visualized the internal structure of calcite crystals ($CaCO_3$; trigonal system, space group $R\bar{3}c$) grown in agarose (a linear polysaccharide consisting of alternating 1,3-linked β-D-galactopyranose and 1,4-linked 3,6-anhydro-α-L-galactopyranose) hydrogels.

Calcite, the most abundant biomineral, has interested scientists because of the exquisite control biological organisms exert over its macro- and microscopic structure and the corresponding advanced mechanical properties of these organic/inorganic composite materials (*1*–*9*). Biogenic minerals, including calcite, are often associated with a fibrous, hydrated, organic matrix. For these reasons, we chose to study the growth of calcite crystals in agarose hydrogels. Similar to biogenic calcite, these gel-grown crystals incorporate the agarose polymers, providing a model system with which to study the internal structure of organic/inorganic composites (*10*, *11*).

Calcite crystals were grown by using the ammonium carbonate method in agarose hydrogels (1 w/v %) (*11*, *16*). The gel-grown crystals are faceted with the characteristic rhombohedral morphology of calcite expressed by six $\{10\bar{1}4\}$ faces (Fig. 1, A and B) (*10*, *11*). For STEM investigations, electron-transparent sections of the crystals were prepared by using a dual-beam FIB (fig. S1). Because of the radiation sensitivity of the samples, the final surfaces of the sections were finished by means of a 2-keV low-dose ion-polishing step in order to remove the damaged layer left behind by previous high–kilo–electron volt fast milling procedures (*16*). The resulting ion-thinned area has a wedge shape, giving a usable working thickness ranging from ~20 nm to 1 μm.

After comparing ADF-STEM and TEM images, we selected ADF-STEM for high-resolution imaging and tomography for its more effective suppression of diffraction contrast and the better fulfillment of the projection requirement (the signal is a monotonic function of the projected mass thickness of the sample), which is an essential requirement for directly interpretable and qualitatively reliable tomographic reconstructions. We optimized imaging conditions before the calcite sample was loaded so as to reduce radiation damage (*16*). Tilt-series images (for tomographic reconstruction) were recorded in areas that were 200 to 500 nm thick, where the features of interest (the incorporated organic fibers) are

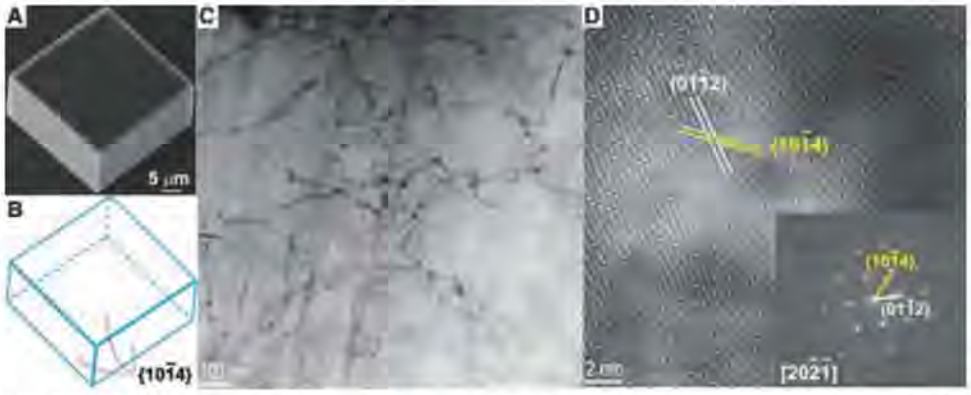


Fig. 1. (**A**) A SEM image of a gel-grown calcite crystal. (**B**) A model of a calcite crystal expressed by six $\{10\bar{1}4\}$ faces. (**C**) A LAADF-STEM image of a thin section cut from a gel-grown calcite crystal by means of FIB. (**D**) A LAADF-STEM lattice image viewed down the $[20\bar{2}\bar{1}]$ zone axis of calcite. (Inset) A SAED pattern of the cut section. The examined area (diameter of 800 nm) contains both crystal and fibers.

[1]Department of Materials Science and Engineering, Cornell University, Ithaca, NY 14853, USA. [2]Department of Physics, Cornell University, Ithaca, NY 14853, USA. [3]School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: lae37@cornell.edu

abundantly sampled and the imaging quality is not substantially degraded by beam divergence and beam spreading (*17*).

Low-angle ADF-STEM (LAADF-STEM) images of a thin section of gel-grown calcite reveal an interconnected network of darker fibers (average diameter of 13 ± 5 nm) within a brighter matrix (Fig. 1C). These fibrous structures are assigned to be the agarose fibers on the basis of the difference in the elastic scattering cross-section of the two materials. The lower average atomic number (Z) of the organic polymer as compared with that of the calcite results in the observed difference in Z-contrast between the two materials.

In order to examine the 3D structure of the polymer network, tomographic reconstructions were generated from a tilt series of high-angle ADF-STEM (HAADF-STEM) images (*16*, *18*).

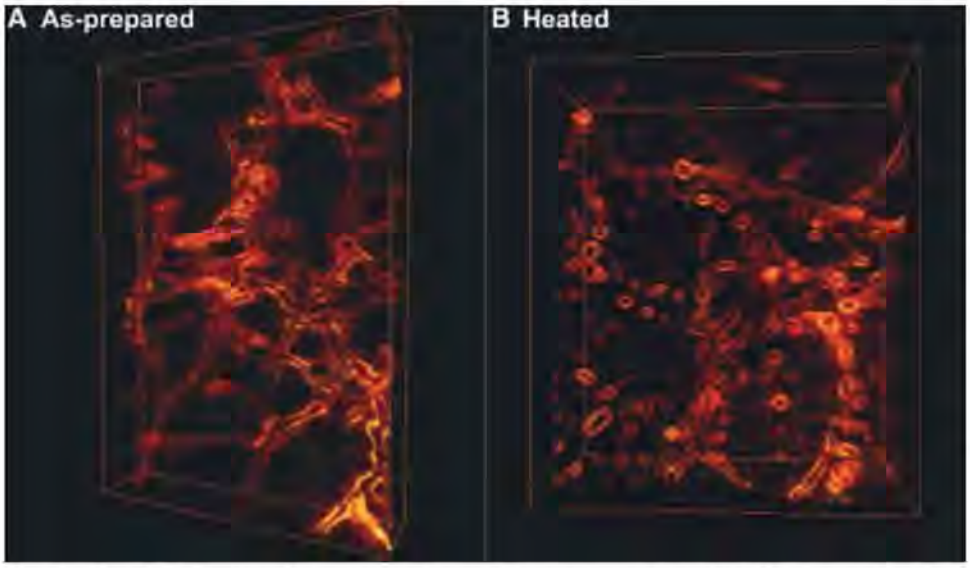


Fig. 2. (**A** and **B**) Tomographic reconstructions of (A) an agarose network inside of a section of as-prepared crystal and (B) cavities inside of a section of heated crystal. The fiber/crystal and cavity/crystal interfaces were highlighted (made brighter) by using a 3D Sobel filter for enhanced contrast. A Sobel filter takes the local gradient of the image, removing any slowly varying background (*16*). The fiber networks are already visible in both the density rendered images (fig. S2) and the raw tilt series (movies S3 and S4). The view is reconstructed from a tilt series of HAADF-STEM images recorded from approximately –70° to 70° [–68° to 60° for (A) and –55° to 61° for (B)] at 2° intervals. Movies of the rotating views showing the 3D structures are provided in the supporting online material (movies S1 and S2). The dimensions of the bounding box of the 3D reconstructions are (A) 1453 nm by 975 nm by 220 nm and (B) 1405 nm by 1196 nm by 554 nm.

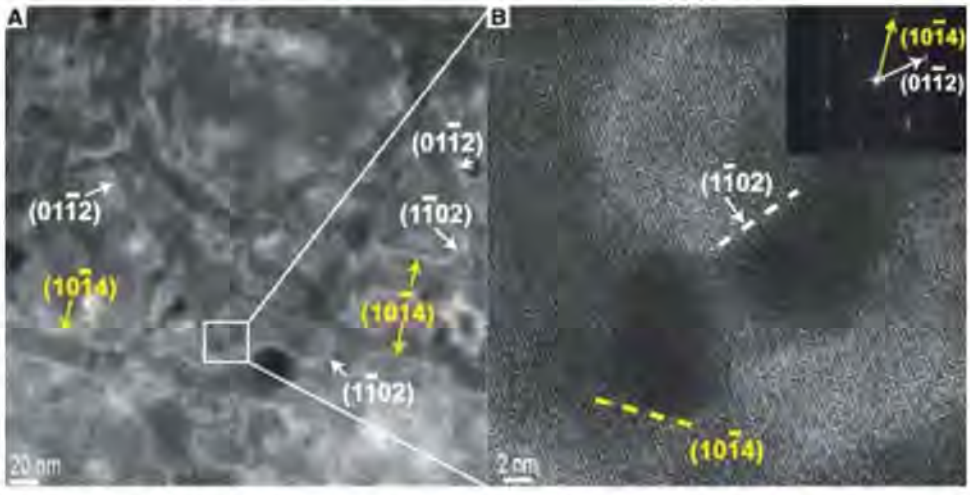


Fig. 3. (**A** and **B**) LAADF-STEM images of a thin section cut from a gel-grown calcite crystal by means of FIB. In (B), the lattice fringes are continuous across the fiber because there is a considerable volume of crystal above and/or below the fiber. (Inset) Fast Fourier transform (FFT) of (B) [a magnified view of (B) showing lattice fringes is provided as fig. S4]. The vertical streaking in the FFT is due to scan noise (explained in detail in fig. S5). On the basis of the FFT, interfaces between the crystal and fibers are partially indexed. For clarity, faces in the {10$\bar{1}$4} family are indicated in yellow, whereas faces in the {01$\bar{1}$2} family are highlighted in white.

The result shows that the incorporated polymer aggregates form a 3D random network penetrating throughout the section of the calcite crystal (Fig. 2A, fig. S2, A and B, and movies S1 and S3). The preservation of the fiber network indicates that the agarose fibers are relatively rigid and that calcite grows around the fibers without substantially disrupting the fiber aggregates (*19*). Although the examined section was lifted out from near the surface of a gel-grown crystal (~10 μm penetration of ~60 μm total height), we believe for two reasons that the images are representative of the interior of the whole crystal. First, the network persists over our largest field of view (~4 μm by 4 μm) (fig. S3). Second, as we have reported previously, gentle etching of fractured crystals revealed that polymer fibers are distributed throughout the crystal (*19*).

Despite the presence of the organic networks, the crystals maintain their single-crystal nature. Selected-area electron diffraction (SAED) of a large area (diameter of 800 nm), including fibers, gives a single set of diffraction spots (Fig. 1D, inset). In addition, high-resolution LAADF-STEM images directly show the regular 2D lattice image viewed down the [20$\bar{2}$1] zone axis of calcite (Fig. 1D).

The incorporation of polymer fibers into single crystals of calcite introduces interfaces between the organic fibers and the inorganic crystals. Indexing the facets that form the internal interfaces shows that the interfaces are a mixture of {10$\bar{1}$4} low-energy facets and {01$\bar{1}$2} high-energy (homocharged) facets (Fig. 3 and fig. S6). In addition to these facets, there are other high-energy facets that could not be indexed. If the sample had been imaged at varied orientations (zones), very possibly we would observe additional high-energy facets forming the crystal/polymer interfaces.

At the length scale of hundreds of nanometers (Figs. 1C and 2A), the interfaces are curved and thus are defined by the contours of the flexible polymer fibers. Because the fibers are interconnected as 3D random networks, at large length scales the interfaces are 3D random curved surfaces. At smaller length scales (tens of nanometers), however, the interfaces appear to be faceted and thus are defined by the faceted habit of the crystal (Fig. 3). A combination of high- and low-energy facets is observed surrounding the fibers. Facets in addition to the {10$\bar{1}$4} facets are required to satisfy the local curvature of the fibers and to minimize the interface area.

There are several possible mechanisms with which to explain the appearance of high-energy facets surrounding the agarose fibers. We will consider two general categories: physical and chemical mechanisms. Chemical (or epitaxial) mechanisms, usually involving molecular recognition between an organic additive and a specific face of calcite, are commonly invoked to explain the expression of non-{10$\bar{1}$4} surfaces in calcite crystals (*20*–*26*). For example, {01$\bar{1}$2} faces are seldom expressed, except in the pres-

ence of homocharged (anionic) surfaces, which are believed to template the formation of a layer of calcium ions with the same structure as the $\{01\bar{1}2\}$ faces (*27*–*29*). It is possible that the agarose fibers may in a similar way stabilize or template the formation of the $\{01\bar{1}2\}$ faces. The external morphology of the gel-grown calcite crystals (six $\{10\bar{1}4\}$ faces), however, suggests that the uncharged polysaccharide, agarose, has a weak affinity for calcite (Fig. 1, A and B). Furthermore, even though both the external and internal surfaces are exposed to the agarose polymers, high-energy facets only appear at the internal surfaces.

As an alternative or supplement to a chemical mechanism, we propose that the coexistence of singular $\{10\bar{1}4\}$ faces at the external surfaces and the nonsingular faces at the internal surfaces (Fig. 3) possibly originates from the different geometry (convex versus concave) of the growth fronts. According to the theory of crystal growth, the external morphology of crystals is defined by slow-growing (usually low-energy) facets, as depicted by kinetic Wulff diagrams. The external facets, by definition, are convex growth fronts. The picture changes for internal surfaces, which are concave. It is predicted that such concave growth fronts will instead be dominated by fast-growing (usually high-energy) faces (fig. S7) (*30*). The external surfaces of calcite crystals are convex; therefore, the low-energy (usually slow-growing) $\{10\bar{1}4\}$ faces dominate the final morphology (Fig. 1, A and B, and fig. S7B). In contrast, the internal growth fronts are concave because the crystals must grow around the fibers to incorporate them, satisfying the curvature of the fibers (fig. S7A). Consequently, we would predict the appearance of multiple, high-energy facets at the internal interfaces, which is consistent with the observation of the high-energy (usually fast-growing) faces, including $\{01\bar{1}2\}$ faces at the calcite/agarose interfaces (Fig. 3 and fig. S7, A and C). As described earlier, although $\{01\bar{1}2\}$ faces are exemplified as high-energy (fast-growing) facets there are other high-energy facets present. Although this physical model is consistent with our observations, further work is required to determine the formation mechanism (or mechanisms) of the high-energy facets, including possible contributions from molecular recognition processes, which might favor one set of facets over another.

In order to investigate the stability of the internal porous structure, we removed the organic network by heating the crystals at 400°C under a flowing air atmosphere for one hour (*16*). The removal of agarose fibers at this temperature was confirmed through thermogravimetric analysis (fig. S8) (*19*). Detailed ADF-STEM observations show the differences between the internal structure of the crystals before and after heating (Fig. 4). After pyrolysis, cavities are observed instead of continuous fiber-like interfaces, as highlighted by the arrows in Fig. 4A. Tomographic reconstructions show in three dimensions that these cavities often are arranged in lines (Fig. 2B, fig. S2, C and D, and movies S2 and S4). This observation suggests that the original, continuous channels fragmented into discrete cavities after the polymer fibers were removed. This surface evolution could be due to reorganization of the internal surfaces to reduce surface area. ADF-STEM images of individual cavities reveal that the high-energy $\{01\bar{1}2\}$ facets are still present at the internal surfaces after removal of the organic material (Fig. 4, B to D). SAED and LAADF-STEM lattice images show that even after pyrolysis, the crystals behave as single crystals (Fig. 4, E and F). Therefore, crystalline integrity and high-energy internal facets are both preserved after pyrolysis. The internal surfaces, however, become isolated and disconnected from the external surfaces and are no longer accessible.

In summary, we have identified the mode of distribution of agarose fibers inside calcite single crystals and observed the polymer/crystal interfaces at high resolution. In these synthetic single crystals, the incorporated macromolecular aggregates are distributed as a 3D network of nanofibers as opposed to isolated molecules or individual fibers. The internal, curved polymer/crystal interfaces are formed by a combination of high- and low-energy facets. These results support physical entrapment of polymer aggregates (*19*) as a possible mechanism by which macromolecules can become incorporated inside inorganic single-crystals. In contrast to other mechanisms, which invoke a specific chemical interaction between the organic component and

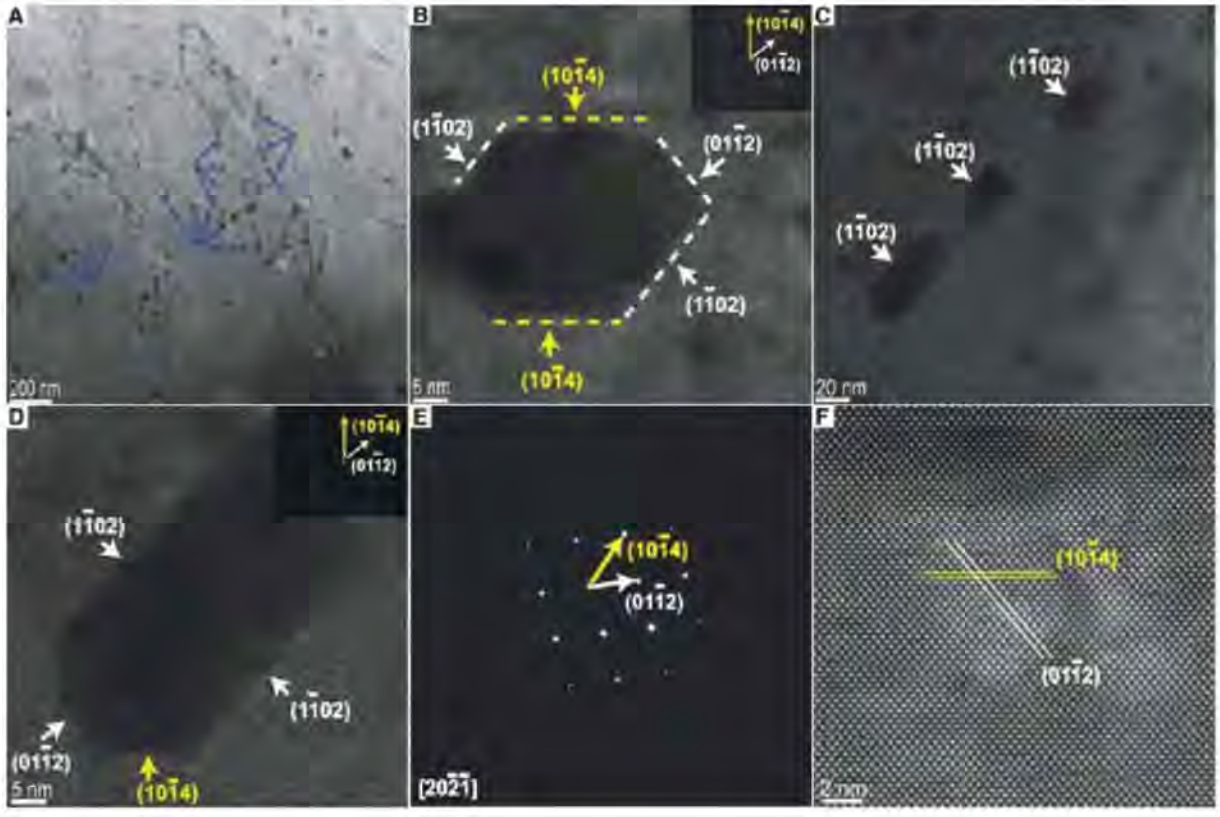


Fig. 4. (**A** to **D**) ADF-STEM images of a thin section (prepared by means of FIB) cut from a gel-grown calcite crystal after heating. The (B) inset shows an FFT of (B), and the (D) inset shows an FFT of (D). On the basis of the FFTs, the interfaces between the crystal and cavities are partially indexed. For clarity, faces in the $\{10\bar{1}4\}$ family are indicated in yellow, whereas faces in the $\{01\bar{1}2\}$ family are highlighted in white. (A) and (B) were taken in LAADF imaging mode in order to give a better signal-to-noise ratio and higher contrast between the calcite and the cavities in thin regions of the section. (C) and (D) were taken in HAADF imaging mode in order to eliminate diffraction contrast, allowing the calcite/cavity interfaces to be imaged sharply in thicker regions of the sample. (**E**) A SAED pattern of the section. The examined area (diameter of 800 nm) contains both crystals and internal cavities. (**F**) A LAADF-STEM image showing the lattice of a heated calcite crystal viewed along the $[20\bar{2}1]$ zone axis.

the growing crystal (*4*, *15*), physical entrapment can occur for a wide range of organic material. This work suggests an approach for modifying the internal structures of crystals and synthesizing single-crystal composites with large, potentially accessible, internal surface areas. Potential uses for the gel method include the preparation of materials that require both high crystallinity and high surface areas, such as photovoltaic materials (*31*).

References and Notes

1. Y. Politi, T. Arad, E. Klein, S. Weiner, L. Addadi, *Science* **306**, 1161 (2004).
2. J. Aizenberg, D. A. Muller, J. L. Grazul, D. R. Hamann, *Science* **299**, 1205 (2003).
3. J. Aizenberg, A. Tkachenko, S. Weiner, L. Addadi, G. Hendler, *Nature* **412**, 819 (2001).
4. J. Aizenberg, J. Hanson, T. F. Koetzle, S. Weiner, L. Addadi, *J. Am. Chem. Soc.* **119**, 881 (1997).
5. J. S. Robach, S. R. Stock, A. Veis, *J. Struct. Biol.* **151**, 18 (2005).
6. F. Nudelman, H. H. Chen, H. A. Goldberg, S. Weiner, L. Addadi, *Faraday Discuss.* **136**, 9 (2007).
7. A. Berman *et al.*, *Science* **259**, 776 (1993).
8. A. Berman *et al.*, *Science* **250**, 664 (1990).
9. B. Pokroy *et al.*, *J. Struct. Biol.* **155**, 96 (2006).
10. H. Y. Li, L. A. Estroff, *J. Am. Chem. Soc.* **129**, 5480 (2007).
11. H. Y. Li, L. A. Estroff, *CrystEngComm* **9**, 1153 (2007).
12. S. Sindhu *et al.*, *Adv. Funct. Mater.* **17**, 1698 (2007).
13. R. J. Park, F. C. Meldrum, *Adv. Mater.* **14**, 1167 (2002).
14. C. Li, L. M. Qi, *Angew. Chem. Int. Ed.* **47**, 2388 (2008).
15. A. N. Kulak *et al.*, *J. Am. Chem. Soc.* **129**, 3729 (2007).
16. Materials and methods are available as supporting material on *Science* Online.
17. J. K. Hyun, P. Ercius, D. A. Muller, *Ultramicroscopy* **109**, 1 (2008).
18. P. A. Midgley, M. Weyland, *Ultramicroscopy* **96**, 413 (2003).
19. H. Y. Li, L. A. Estroff, *Adv. Mater.* **21**, 470 (2009).
20. J. J. De Yoreo, P. M. Dove, *Science* **306**, 1301 (2004).
21. C. A. Orme *et al.*, *Nature* **411**, 775 (2001).
22. S. Mann *et al.*, *Science* **261**, 1286 (1993).
23. L. A. Estroff, C. D. Incarvito, A. D. Hamilton, *J. Am. Chem. Soc.* **126**, 2 (2004).
24. N. Sommerdijk, G. de With, *Chem. Rev.* **108**, 4499 (2008).
25. E. M. Pouget *et al.*, *Science* **323**, 1455 (2009).
26. R. Q. Song, A. W. Xu, M. Antonietti, H. Colfen, *Angew. Chem. Int. Ed.* **48**, 395 (2009).
27. J. Aizenberg, A. J. Black, G. M. Whitesides, *Nature* **398**, 495 (1999).
28. A. M. Travaille *et al.*, *J. Am. Chem. Soc.* **125**, 11571 (2003).
29. D. Volkmer, M. Fricke, C. Agena, J. Mattay, *J. Mater. Chem.* **14**, 2249 (2004).
30. D. X. Du, D. J. Srolovitz, M. E. Coltrin, C. C. Mitchell, *Phys. Rev. Lett.* **95**, 155503 (2005).
31. I. Gur, N. A. Fromer, M. L. Geier, A. P. Alivisatos, *Science* **310**, 462 (2005).

32. L.A.E. acknowledges partial support from NSF (DMR 0845212), the J. D. Watson Investigator Program (New York State Office of Science, Technology, and Academic Research contract C050017), and the Cornell Center for Materials Research (CCMR), a Materials Research Science and Engineering Center of NSF (DMR 0520404). H.L.X. and D.M. acknowledge the Semiconductor Research Corporation and CCMR. Particular acknowledgement is made of the use of the EM and polymer characterization facilities of CCMR.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1244/DC1
Materials and Methods
Figs. S1 to S8
References
Movies S1 to S4

2 July 2009; accepted 14 October 2009
10.1126/science.1178583

Formation of Compositionally Abrupt Axial Heterojunctions in Silicon-Germanium Nanowires

C.-Y. Wen,[1] M. C. Reuter,[2] J. Bruley,[2] J. Tersoff,[2] S. Kodambaka,[3] E. A. Stach,[1] F. M. Ross[2]*

We have formed compositionally abrupt interfaces in silicon-germanium (Si-Ge) and Si-SiGe heterostructure nanowires by using solid aluminum-gold alloy catalyst particles rather than the conventional liquid semiconductor–metal eutectic droplets. We demonstrated single interfaces that are defect-free and close to atomically abrupt, as well as quantum dots (i.e., Ge layers tens of atomic planes thick) embedded within Si wires. Real-time imaging of growth kinetics reveals that a low solubility of Si and Ge in the solid particle accounts for the interfacial abruptness. Solid catalysts that can form functional group IV nanowire-based structures may yield an extended range of electronic applications.

The most exciting applications of semiconductor nanowires—for electronic devices as well as for fundamental research into transport phenomena—do not come from compositionally uniform structures, but from heterostructure nanowires in which the composition changes along the length of the wire. These axial heterostructures are particularly important for applications such as tunnel field-effect transistors and thermoelectric devices (*1*–*4*). A prerequisite for optimal device performance in these applications is the formation of compositionally abrupt and structurally perfect junctions (*5*, *6*). The relative ease of forming abrupt, defect-free junctions in group III-V nanowires composed of materials pairs such as InAs-InP has been key to realizing devices such as double-barrier resonant tunneling devices and single-electron transistors (*3*, *7*–*9*). Yet abrupt interfaces have not been formed in group IV nanowires, creating a severe constraint on the use of group IV materials in nanowire applications. This is thought to be due to the fundamental nature of the process by which the nanowires grow.

Si and Ge nanowires are commonly fabricated using the vapor-liquid-solid (VLS) mechanism (*10*). In this method, droplets of a liquid catalyst are formed on a substrate via eutectic reaction between the semiconductor and a metal, typically gold, above the eutectic temperature T_{eu}. A vapor-phase semiconductor reactant, such as silane or germane, is then introduced and is preferentially dissociated by the liquid catalyst and dissolved into the liquid. Once the concentration of semiconductor atoms in the catalyst reaches supersaturation, the semiconductor precipitates at the liquid-substrate interface. Nanowires then grow as material is added in a layer-by-layer fashion at the liquid-solid interface. To create a junction between two different semiconductors, the vapor-phase reactants are modulated during growth. This, of course, requires a catalyst suitable for the growth of both components under compatible conditions.

Group IV heterostructure nanowires containing dislocation-free Si and SiGe segments have been grown successfully using this approach (*11*, *12*). However, the interfaces are diffuse, showing a broad composition gradient. A similar broadening of Si-Ge interfaces has also been seen for "MBE-grown" nanowires, where the same Au starting material is used but the Si and Ge are supplied by molecular beam epitaxy (*13*, *14*). Theoretical modeling and experimental investigation (*12*, *15*) suggest that this interface diffuseness is due to a "reservoir effect." The catalytic droplet contains a large amount of the growth species; AuSi and AuGe contain more than 20 atomic percent of Si or Ge, respectively. This means that when the reactant supply is switched, say from silane to germane, the droplet must be depleted of Si before pure Ge can be deposited. It is found (*12*) that the length of the compositionally graded region scales with wire diameter, as expected from the modeling (*12*, *15*). Because interface diffuseness is due to the high solubility of Si and Ge in the catalyst, it cannot be avoided with the usual liquid semiconductor–metal catalysts. However, in solid metal catalysts the solubility of Si or Ge can be very low, so that rapid depletion of each material would be expected. Our approach to an abrupt interface in a nanowire junction structure exploits the potential of solid catalysts and uses the vapor-solid-solid (VSS) growth mechanism.

[1]School of Materials Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907, USA. [2]IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA. [3]Department of Materials Science and Engineering, University of California, Los Angeles, CA 90095, USA.

*To whom correspondence should be addressed. E-mail: fmross@us.ibm.com

Several solid catalysts have been used to grow Si and Ge nanowires (*16*–*22*). We ruled out the use of metal silicides or germanides (*18*–*22*) for our purpose, because of possible reservoir effects due to transformations between silicide and germanide during formation of the heterostructure. Solid Au catalyzes Ge wire growth (*16*), but because the reaction must be carried out below T_{eu} (~360°C for both Si and Ge), VSS growth is impractically slow for conventional Si and Ge reactants. Aluminum, which is of interest for photovoltaic nanowire applications because of its electronic properties in Si (*23*), has a higher T_{eu} (577°C) but slow Si wire growth kinetics (*17*, *24*). To realize efficient heterojunction nanowire growth, we sought a solid catalyst that is suitable for both Si and Ge and that can provide practical growth rates. To this end, we added Al to Au catalysts (i) to raise the eutectic temperature with Si or Ge, and hence the nanowire growth temperature and growth rate under VSS, while (ii) maintaining the high reactivity known for Au. The ternary phase diagrams of Al-Au-Si and Al-Au-Ge have not been determined experimentally, but theoretical calculations predict a eutectic temperature of Si with one of the intermetallic compounds, $AlAu_2$, of 487°C (*25*). Because of the similarity in physical properties between Si and Ge, we might expect Ge to have a similar T_{eu} with $AlAu_2$.

Growth experiments were conducted in a transmission electron microscope (TEM) (*26*), because real-time observations proved useful in providing structural information during growth as well as in tuning the growth parameters (*27*). The catalyst was prepared by thermally evaporating Au and Al sequentially with an atomic ratio of 2:1. Two growth modes are possible in this system, the faster VLS and the slower VSS, with T_{eu} around the expected value (Fig. 1, A to C). Figure 1, D to I, shows the growth sequence of a Si-Ge-Si heterostructure as it forms in a single nanowire. We first grew long Si nanowires in the VLS mode and then cooled them below T_{eu} to obtain the VSS mode (Fig. 1D). Subsequent modulation in the gas source produced the junction segments effectively (Fig. 1, E to I). Under these imaging conditions, Ge exhibits darker contrast, and the sharp change between the narrow Ge layer (or "quantum dot") and the surrounding Si hints at compositionally abrupt interfaces. The VSS growth rates for Ge and Si are slower than the rapid VLS Si nanowire growth but are still practical for the purpose of growing the heterostructure.

We quantify the structural perfection and the compositional abruptness of the Si-Ge interface in Fig. 2, where we have simplified the structure by growing only a Ge layer on the tip of a Si nanowire in order to avoid ambiguity in analysis due to overlay of materials on wire sidewalls. The analysis in Fig. 2 was obtained after (rather than during) growth, allowing the use of microscopes with analytical capabilities and higher spatial resolution but requiring the sample to be exposed to air (*26*). In Fig. 2A we see that the interface exhibits a single crystalline structure without obvious crystal defects such as dislocations. High-angle annular dark-field scanning TEM (HAADF-STEM; Fig. 2B) confirms the

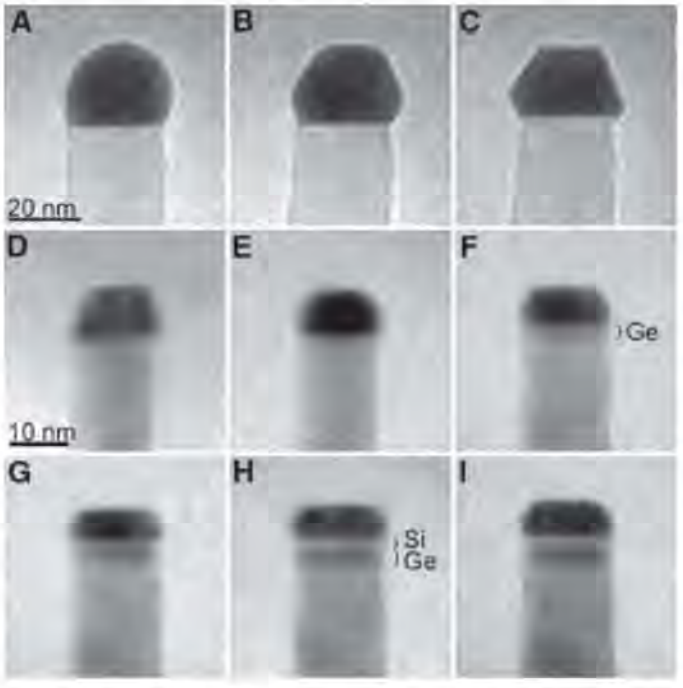


Fig. 1. (**A** to **C**) Frames extracted from a movie recorded during growth of a Si nanowire (diameter 28 nm) in 3×10^{-5} torr disilane as the temperature was reduced from 570°C to 503°C: (A) immediately before solidification of the catalyst (507°C); (B) just after solidification (503°C), which is evident from the faceted shape of the catalyst; (C) 14 s after (B), growth rate 5 nm/min. The growth rate is 22 nm/min in the VLS mode at 570°C. (**D** to **I**) Formation of a Si-Ge-Si heterojunction in a Si nanowire (diameter 14 nm). The Si nanowire in (D) was first grown at 510°C with 1×10^{-5} torr Si_2H_6 for 2 hours and then cooled to 360°C. The faceted surface after cooling indicates the solid state of the catalyst. (**E**) After flowing 5×10^{-6} torr Ge_2H_6 at 360°C for 1 min, Ge growth has begun and the catalyst has become less sharply faceted. (**F**) After 7 min, formation of a thin Ge layer that appears as a dark band at the Si-catalyst interface. (**G**) After 8.5 min of Si growth at 360°C and 1×10^{-5} torr Si_2H_6. (**H**) After 14 min of Si growth, the catalyst has resumed the strong facets on its surface seen in (D). (**I**) For demonstration purposes, the catalyst was exposed to 5×10^{-6} torr Ge_2H_6 at 360°C for 1 min again, resuming the shape seen in (F). The growth rates of the Ge and Si segments measured in (H) are 0.4 and 0.2 nm/min, respectively. The growth direction of the wires is Si[111]. The dark contrast on the Si segment in (G) to (I) is due to conformal deposition of Ge on the Si sidewalls.

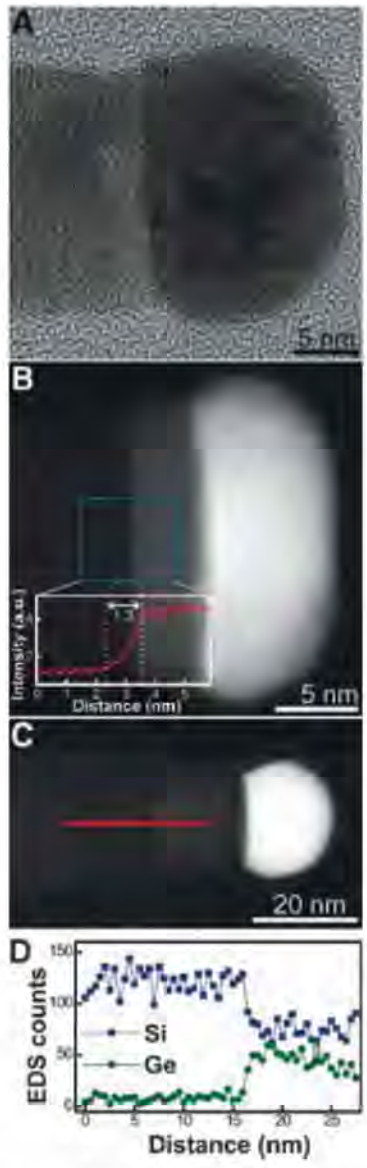


Fig. 2. (**A**) High-resolution TEM image of a Si-Ge heterojunction nanowire. A Si wire (length 500 nm, diameter 18 nm) was grown rapidly using VLS at 510°C with 1×10^{-5} torr Si_2H_6 for 2 hours, then cooled and grown for several minutes in the VSS mode at 360°C, followed by growth of a Ge segment at 360°C with 5×10^{-6} torr Ge_2H_6 for 16 min. (**B**) HAADF-STEM image of a wire (diameter 17 nm) grown under the same conditions. The inset shows the intensity profile across the interface, averaged over a 5-nm strip along the midpoint of the wire. The width of the interface is 1.3 nm. (**C**) HAADF-STEM image of a $Si/Si_{1-x}Ge_x$ nanowire (diameter 21 nm). Si was grown at 510°C and 1×10^{-5} torr Si_2H_6 for 2 hours; $Si_{1-x}Ge_x$ was grown at 430°C with 2×10^{-6} torr Si_2H_6 and 1×10^{-7} torr 20% Ge_2H_6; then Si was grown at 510°C and 2×10^{-6} torr Si_2H_6. (**D**) EDS line profile of Si and Ge through the $Si/Si_{1-x}Ge_x$ junction, as indicated in (C), showing a sharp transition (less than 2 nm) from Si to SiGe. The composition of the $Si_{1-x}Ge_x$ alloy segment is estimated to be $Si_{0.7}Ge_{0.3}$.

presence of a compositionally uniform Ge segment on the Si nanowire. The average intensity profile (Fig. 2B, inset) shows a rather narrow transition (~1.3 nm) at the interface. STEM energy dispersive x-ray spectroscopy (EDS) line profiles (*26*) (Fig. 2D and fig. S1) show that the composition transition occurs in less than 2 nm, in agreement with the HAADF analysis. EDS maps (Fig. 3) confirm the uniform Ge distribution in the Ge layer; they show that there is no Si or Ge in the bulk of the catalyst and no Si in the Ge (at the detection limit, ~2%). These results can be compared with the diffuseness of interfaces measured (using similar techniques) during a comprehensive study of Si-SiGe heterostructures grown using liquid catalysts (*12*); wires of the diameter shown in Fig. 2 were found to have diffuse interfaces with compositional widths on the order of 18 to 24 nm.

Apart from the reservoir effect, other factors may affect the measured interface abruptness, including Si-Ge interdiffusion at the growth temperature and broadening of the STEM probe. A previous study of Ge self-diffusion in $Si_{1-x}Ge_x$ ($0 \leq x \leq 0.5$) under various strain conditions suggested that Si-Ge interdiffusion at the nanowire growth temperature should be slow enough to be negligible over our growth times (*28*). Electron beam broadening, however, does cause a noticeable error in determining interface abruptness (*29*). Using Goldstein's single-scattering equation (*29*), we estimate a beam broadening of 0.3 nm in Si and 0.6 nm in Ge (using a wire thickness of 20 nm and an electron energy of 300 keV). Adding these effects to the original 0.5-nm probe, an atomically abrupt interface should appear with a width of ~1 nm. In other words, the compositional width of our Si-Ge interface is probably less than that shown in the profiles in Figs. 2 and 3 and fig. S1, and may be below 1 nm. Any physical roughness (nonplanarity of the interface) would also contribute to the measured interface abruptness. We cannot completely rule out the possibility that some Si could remain in the catalyst, but any reservoir effect must be small, as we do not see a measurable difference between the widths of the interfaces in wires 20 nm and 10 nm in diameter (*26*) (fig. S1). Because Ge and Si are not detectable in the catalyst, their solubility is low, fully supporting the advantage of using a solid catalyst to reduce the reservoir effect.

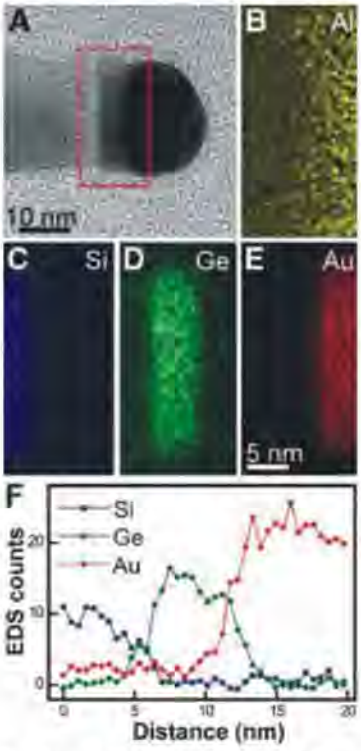


Fig. 3. Composition analysis of a Si-Ge heterojunction nanowire grown under the conditions of Fig. 2A. (**A** to **E**) Bright-field TEM image (A) and false-color STEM EDS elemental maps (B to E) of Al, Si, Ge, and Au in the region [defined in (A)] of the junction in a nanowire 22 nm in diameter. The catalyst is surrounded by an amorphous shell, 2 nm thick, that forms after air exposure. The stronger Al signal in this region and the difference in interface diffuseness between the Al and Au maps suggests that the shell is aluminum oxide, consistent with the results of oxidation of AlAu alloys described in (*30*). (**F**) Line profile of the EDS intensities extracted from the elemental maps of Si, Ge, and Au. The intensity is averaged over a 3-nm strip along the midpoint of the wire.

During VSS growth, ledge nucleation and flow at the catalyst-wire interface is evident (fig. S2) and is expected in VSS systems (*20*). Here, by measuring the kinetics of individual ledge motion, we found evidence to support a solubility of Si and Ge in the solid catalyst that is low relative to the solubility in liquid Au eutectic droplets. During VSS growth of Si, we found that ledges nucleate only a short time after a previous ledge has finished propagating (fig. S2). In contrast, for VLS growth at the same overall growth rates, we observed quite different ledge flow kinetics (fig. S3): long "incubation times" (several seconds) between ledges, followed by very rapid ledge propagation at a rate too fast to measure. In VLS, it is not surprising that substantial excess Si or Ge is needed to raise the chemical potential high enough to nucleate a ledge. This excess provides a large driving force and supply for the subsequent rapid motion. For VSS, the short incubation time suggests that even a tiny amount of excess Si raises the chemical potential enough to nucleate a ledge, consistent with low solubility. In that case, the excess Si would be too small to drive the ledge very far, and subsequent motion would then be supply-limited, depending directly on the incoming flux.

The fact that ledge motion depends directly on the incoming flux is particularly relevant to the growth of SiGe alloy segments in nanowires. When growing alloy segments from liquid eutectic droplets, the composition profiles at Si-SiGe interfaces are found to be curved (*12*); in other words, the central region of the wire switches composition at a different rate relative to regions nearer the wire surface. This effect is thought to arise because, although the growth interface remains planar, the incorporation ratio of Si and Ge changes as a result of varying strain (*12*). Here, we expect such effects to be avoided or minimized because of the direct dependence of the deposited composition on the gas supply. Indeed, SiGe segments grown from solid catalysts (Fig. 2, C and D) by flowing disilane and digermane simultaneously show abrupt and planar Si-SiGe interfaces (<2 nm wide) and a uniform composition within the SiGe segment.

The following are some practical considerations for extending the technique of using a solid catalyst with low group IV solubility to create abrupt heterointerfaces in nanowires.

1) We have mentioned the need for a relatively high eutectic temperature, so that the VSS growth mode can be obtained at temperatures high enough to achieve reasonably fast growth rates. However, it is also useful to be able to access the VLS growth mode when needed, by means of a small increase in temperature. This was shown in Fig. 1 where the VLS mode was used to expedite Si nanowire growth, forming wire segments hundreds of nanometers in length, after which the sample was cooled to obtain VSS Si growth before forming Si-Ge interfaces or quantum dots.

2) To form wires consistently with a particular morphology (i.e., a well-defined cross-sectional shape and growth direction and no kinking), we suggest that it is advantageous to have a strong orientation relationship between the solid catalyst crystal and the wire. For the majority of wires, the $AlAu_2$ catalyst we have used here exhibits a large top facet and a flat shape, and this orientation forms wires that are generally unkinked.

3) The solid catalyst must be able to grow both Si and Ge segments without changes in its shape, because shape changes could alter the wire diameter.

It is our belief that other solid materials that catalyze Si and Ge wire growth should be useful in forming abrupt heterointerfaces, provided that they satisfy the requirements above in terms of eutectic temperature, the existence of a preferred orientation relationship, and minimal changes in shape between growing Si and growing Ge.

Our results show that axial Si-Ge nanowire heterostructures in which the composition is modulated at the nanometer scale can be realized by using an AlAu alloy catalyst operating via the VSS growth mechanism. This method provides control over compositional modulation that is superior to that possible with the use of conventional AuSi-AuGe liquid catalysts. The key to achieving this control is twofold: The solid state of the catalyst leads to a low solubility of Si or Ge and avoids the reservoir effect that accounts for the smearing of compositional changes, and the tuning of the eutectic temperature allows rapid growth with a VLS mode to form long uniform segments, as well as VSS growth at a slower yet practical rate for nanoscale control of interfaces,

quantum dots, and barriers. We believe that this approach should apply to solid catalyst materials other than AlAu alloys, and indeed to wire materials other than Si and Ge, provided that the catalyst and wire material satisfy requirements for eutectic temperature, orientation, and stability. We speculate that a solid catalyst could be similarly advantageous in forming abrupt doping profiles, such as are required for high-subthreshold slope devices such as tunnel and avalanche field-effect transistors. More generally, Si and Ge device fabrication that requires sophisticated control of composition and structure can be addressed using this approach.

References and Notes

1. M. S. Gudiksen, L. J. Lauhon, J. Wang, D. C. Smith, C. M. Lieber, *Nature* **415**, 617 (2002).
2. C. Yang, Z. Zhong, C. M. Lieber, *Science* **310**, 1304 (2005).
3. M. T. Björk *et al.*, *Nano Lett.* **4**, 1621 (2004).
4. D. Li, Y. Wu, R. Fan, P. Yang, A. Majumdar, *Appl. Phys. Lett.* **83**, 3186 (2003).
5. C. M. Lieber, *Nano Lett.* **2**, 81 (2002).
6. L. Samuelson *et al.*, *Physica E* **25**, 313 (2004).
7. M. T. Björk *et al.*, *Appl. Phys. Lett.* **81**, 4458 (2002).
8. M. T. Björk *et al.*, *Nano Lett.* **2**, 87 (2002).
9. C. Thelander *et al.*, *Appl. Phys. Lett.* **83**, 2052 (2003).
10. R. S. Wagner, W. C. Ellis, *Appl. Phys. Lett.* **4**, 89 (1964).
11. Y. Wu, R. Fan, P. Yang, *Nano Lett.* **2**, 83 (2002).
12. T. E. Clark *et al.*, *Nano Lett.* **8**, 1246 (2008).
13. N. D. Zakharov *et al.*, *J. Cryst. Growth* **290**, 6 (2006).
14. R. Dujardin *et al.*, *Appl. Phys. Lett.* **89**, 153129 (2006).
15. N. Li, T. Y. Tan, U. Gösele, *Appl. Phys. A* **90**, 591 (2008).
16. S. Kodambaka, J. Tersoff, M. C. Reuter, F. M. Ross, *Science* **316**, 729 (2007).
17. Y. Wang, V. Schmidt, S. Senz, U. Gösele, *Nat. Nanotechnol.* **1**, 186 (2006).
18. T. I. Kamins, R. S. Williams, D. P. Basile, T. Hesjedal, J. S. Harris, *J. Appl. Phys.* **89**, 1008 (2001).
19. J. Arbiol, B. Kalache, P. Roca i Cabarrocas, J. R. Morante, A. Fontcuberta i Morral, *Nanotechnology* **18**, 305606 (2007).
20. S. Hofmann *et al.*, *Nat. Mater.* **7**, 372 (2008).
21. K. Kang *et al.*, *Adv. Mater.* **20**, 4684 (2008).
22. J. L. Lensch-Falk, E. R. Hemesath, F. J. Lopez, L. J. Lauhon, *J. Am. Chem. Soc.* **129**, 10670 (2007).
23. J. R. Davis *et al.*, *IEEE Trans. Electron. Dev.* **27**, 677 (1980).
24. B. A. Wacaser *et al.*, *Nano Lett.* **9**, 3296 (2009).
25. M. Hoch, *J. Alloy. Comp.* **220**, 27 (1995).
26. See supporting material on *Science* Online.
27. F. M. Ross, J. Tersoff, M. C. Reuter, *Phys. Rev. Lett.* **95**, 146104 (2005).
28. N. R. Zangenberg, J. Lundsgaard Hansen, J. Fage-Pedersen, A. Nylandsted Larsen, *Phys. Rev. Lett.* **87**, 125901 (2001).
29. D. C. Joy, A. D. Romig, J. I. Goldstein, *Principles of Analytical Electron Microscopy* (Plenum, New York, 1986).
30. H. Piao, M. Suominen Fuller, D. Miller, N. S. McIntyre, *Appl. Surf. Sci.* **187**, 266 (2002).

31. We acknowledge A. Ellis and L. Gignac of IBM for their assistance with experimental aspects of this work. Supported by NSF grants DMR-0606395 and DMR-0907483 and by the University of California Energy Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1247/DC1
Materials and Methods
Figs. S1 to S3
References

26 August 2009; accepted 29 September 2009
10.1126/science.1178606

Selective Phenol Hydrogenation to Cyclohexanone Over a Dual Supported Pd–Lewis Acid Catalyst

Huizhen Liu, Tao Jiang,* Buxing Han,* Shuguang Liang, Yinxi Zhou

Cyclohexanone is an industrially important intermediate in the synthesis of materials such as nylon, but preparing it efficiently through direct hydrogenation of phenol is hindered by over-reduction to cyclohexanol. Here we report that a previously unappreciated combination of two common commercial catalysts—nanoparticulate palladium (supported on carbon, alumina, or NaY zeolite) and a Lewis acid such as $AlCl_3$—synergistically promotes this reaction. Conversion exceeding 99.9% was achieved with >99.9% selectivity within 7 hours at 1.0-megapascal hydrogen pressure and 50°C. The reaction was accelerated at higher temperature or in a compressed CO_2 solvent medium. Preliminary kinetic and spectroscopic studies suggest that the Lewis acid sequentially enhances the hydrogenation of phenol to cyclohexanone and then inhibits further hydrogenation of the ketone.

Cyclohexanone is a key raw material in the synthesis of many useful chemical intermediates, such as caprolactam for nylon 6 and adipic acid for nylon 66 (*1*, *2*). The industrial production of cyclohexanone typically involves either the oxidation of cyclohexane (*3*, *4*) or the hydrogenation of phenol. The former route requires high temperature and generates byproducts such as cyclohexanol and organic acids that complicate purification, and the yield of cyclohexanone is usually low. The phenol hydrogenation route is undertaken through either two-step or one-step processes. In the two-step procedure, phenol is first hydrogenated to cyclohexanol, which in turn is dehydrogenated to cyclohexanone at high temperature. The one-step selective hydrogenation of phenol to cyclohexanone is advantageous from an efficiency standpoint, and the reaction can be conducted in either the gas phase or liquid phase. The gas-phase phenol hydrogenation is usually performed in the temperature range of 150° to 300°C over supported Pd catalysts (*5*–*9*), and different supports have been used, including alumina, that may act as Lewis acids (*9*). The gas-phase process can be carried out easily in continuous reactors for higher throughput. Liquid-phase phenol hydrogenation offers cost and energy savings, because the reaction can be performed at relatively low temperatures (*10*–*15*). Many researchers have contributed to this area, and multiple catalysts have been screened, such as Rh/C (*10*), Rh/C nanofiber (*11*), Pd/hydrophilic C (*12*), Ru/poly(*N*-vinyl-2-pyrrolidone) (PVP) (*13*), Pd/Mg and Pd/Fe (*14*), mesoporous Ce-doped Pd (*15*), and Pd/C (*16*). However, the attainment of high selectivity (>95%) at elevated conversion (>80%) with a satisfactory rate is a great challenge (*8*, *17*), because the cyclohexanone product can be further hydrogenated to cyclohexanol under the reaction conditions (*7*, *15*).

Beijing National Laboratory for Molecular Sciences, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190, China.

*To whom correspondence should be addressed. E-mail: Jiangt@iccas.ac.cn (T.J.); Hanbx@iccas.ac.cn (B.H.)

We report here that Pd/C, Pd/Al_2O_3, and Pd/NaY zeolite (*18*) (NaY zeolite is hereafter denoted NaY) catalysts and solid Lewis acids show excellent synergy in the hydrogenation of phenol to cyclohexanone, together substantially enhancing both activity and selectivity. The reaction can be carried out effectively at temperatures as low as 30°C, and >99.9% conversion of phenol is observed with >99.9% selectivity to cyclohexanone. Separation of the product from the catalyst–Lewis acid system is simple, and the catalyst system can be reused directly. This route has great potential for industrial application.

Our experiments (*18*) showed that dichloromethane is the best reaction solvent among several tested (table S2). Table 1 presents the results of phenol hydrogenation under different conditions over Pd/C, Pd/Al_2O_3, and Pd/NaY catalysts with and without $AlCl_3$. The conversion of phenol was very low, and considerable byproduct was produced when only the Pd/C catalyst was used (Table 1, entry 1), and the reaction did not occur at all when only Lewis acid ($AlCl_3$) was used (Table 1, entry 2). When Pd/C and $AlCl_3$ were used at the same time, the reaction proceeded with a selectivity of >99.9% up to complete conversion at 1.0 MPa of H_2 and a temperature at or below 50°C (Table 1, entries 3 to 5). At higher temperature, the reaction reached completion in 1 hour and the selectivity remained >99% (Table 1, entries 9 and 10). With other conditions fixed, an increase in hydrogen pressure (Table 1, entries 6 to 12) shortened the time to completion but slightly reduced the selectivity to cyclohexanone. Other Lewis acids were also effective in promoting the reaction [Table 1, entries 13 to 18; see also fig. S7 and supporting online material (SOM) text]. The activity and selectivity of the reaction using Pd/Al_2O_3 and Pd/NaY (*18*) were also enhanced effectively by $AlCl_3$ (Table 1, entries 19 to 26). The prospects for recycling Pd/C-$ZnCl_2$ were tested (*18*), and the results indicated that the catalyst system could

be reused after simple separation without a decrease in conversion and selectivity (Table 1, entries 16*, 16†, and 16‡). Preliminary kinetics measurements indicated that the reaction exhibits a standard pseudo–first-order rate dependence on phenol concentration (fig. S9 and SOM text).

We also explored the use of CO_2 as a reaction solvent for its characteristic advantages of waste minimization, ease of product separation, and pressure tunability (*19*–*21*). The Pd/C–Lewis acid systems were very efficient catalysts for the reaction in compressed CO_2 at suitable pressures (Table 2, entries 3 to 12). The reaction in CO_2 also follows standard pseudo–first-order kinetics with respect to phenol (fig. S9). The pressure of CO_2 had a significant effect on the apparent reaction rate. At optimized pressures, both the conversion and selectivity were >99.9% (Table 2, entries 3 to 12), and the apparent reaction rate was faster than that in dichloromethane (Table 1, entry 3, versus Table 2, entries 3 to 6; Table 1, entry 13, versus Table 2, entries 7 to 9; Table 1, entry 15, versus Table 2, entries 10 to 12), which may result mainly from the larger diffusion coefficient of the reactants in CO_2 and the higher miscibility of gaseous H_2 with CO_2 than with liquid solvents (*22*). In addition, CO_2 is a weak Lewis acid, which may also enhance the reaction. To further explore the effect of CO_2 pressure on the reaction rate, we observed the phase behavior of the reaction system directly by using a high-pressure view reactor (*21*). A vapor phase (a mixture of H_2 and CO_2) and a solid phase (a mixture of catalyst, Lewis acid, and phenol) were evident at total pressures of 5.0 MPa (Table 2, entry 1) and 6.5 MPa (Table 2, entry 2). The phenol, Pd, and Lewis acid could not come in contact efficiently, and therefore the reaction proceeded slowly. At higher pressures, a liquid phase formed in addition to the vapor phase, due to the liquefaction of CO_2. The catalyst and Lewis acid were dispersed and the reactants dissolved under stirring, allowing the reaction to proceed effectively. Above 7 MPa, however, the apparent reaction rate decreased with increasing pressure (Table 2, entries 3 to 6), which may have resulted from dilution as the volume of the liquid phase increased. Separation of CO_2, product, and catalyst proved quite facile, because the product could be extracted in situ by CO_2, whereas the solid catalyst and Lewis acid remained in the reactor (*18*). At 30°C and a total H_2/CO_2 pressure of 8.0 MPa, the Pd/C-$AlCl_3$ system could be recycled three times with no loss in conversion or selectivity (Table 2, entries 5 and 13 to 15).

Prior research on the hydrogenation of phenol has indicated that this reaction proceeds mainly in a sequential manner (Fig. 1A) (*15*, *23*, *24*). The benzene ring of phenol is first partially hydrogenated to an enol in step 1, which is unstable and isomerizes rapidly to form cyclohexanone (*23*); cyclohexanone can then be further hydrogenated to form cyclohexanol. Cyclohexanone and cyclohexanol were the only reaction products observed in the present study, and under all reaction conditions, selectivity approached 100% at sufficiently low conversion (fig. S9). We conclude from the results in Tables 1 and 2 that under the likely operation of this two-step mechanism, the supported Pd catalyst–Lewis acid systems are very active for the first step but are not active for the second step under the experimental conditions. This is further discussed below.

Table 1. Hydrogenation of phenol under different conditions. Conversions and selectivities were determined by gas chromatography (GC). Reactants and products were identified by GC mass spectrometry (MS) as well as by comparing retention times to respective standards in GC traces. At >99.9% conversion, the minimum time required for reaction completion is given. Cyclohexanone and cyclohexanol were the only reaction products observed. Reaction conditions were as follows: phenol, 1.0 mmol; Pd [5 weight percent (wt %) in Pd/C or Pd/Al_2O_3; 2.5 wt % in Pd/NaY], 5 mol % relative to phenol; Lewis acid, 10 mol % relative to phenol; solvent, 1 ml of dichloromethane. *p*, pressure; *T*, temperature. Dashes indicate that no catalyst or Lewis acid was added. C=O indicates cyclohexanone, and OH denotes cyclohexanol.

Entry	Catalysts	Lewis acids	P_{H_2}(MPa)	*T* (°C)	Time (hours)	Conversion (%)	Selectivity (%)	
							C=O	OH
1	Pd/C	–	1	30	12	12.6	93.8	6.2
2	–	$AlCl_3$	1	30	12	0	0	0
3	Pd/C	$AlCl_3$	1	30	12	>99.9	>99.9	<0.1
4	Pd/C	$AlCl_3$	1	40	11	>99.9	>99.9	<0.1
5	Pd/C	$AlCl_3$	1	50	7	>99.9	>99.9	<0.1
6	Pd/C	$AlCl_3$	1	80	3	>99.9	99.3	0.7
7	Pd/C	$AlCl_3$	2	80	2.5	>99.9	96.8	3.2
8	Pd/C	$AlCl_3$	3	80	2	>99.9	95.1	4.9
9	Pd/C	$AlCl_3$	0.5	100	1	>99.9	99.2	0.8
10	Pd/C	$AlCl_3$	1	100	0.5	>99.9	99.0	1.0
11	Pd/C	$AlCl_3$	2	100	0.3	>99.9	97.3	2.7
12	Pd/C	$AlCl_3$	3	100	0.2	>99.9	96.2	3.8
13	Pd/C	$InCl_3$	1	30	16	>99.9	>99.9	<0.1
14	Pd/C	$InCl_3$	1	100	4	>99.9	99.3	0.7
15	Pd/C	$ZnCl_2$	1	30	20	>99.9	>99.9	<0.1
16	Pd/C	$ZnCl_2$	1	100	6	>99.9	98.7	1.3
16*	Pd/C	$ZnCl_2$	1	100	6	>99.9	98.3	1.7
16†	Pd/C	$ZnCl_2$	1	100	6	>99.9	98.7	1.3
16‡	Pd/C	$ZnCl_2$	1	100	6	>99.9	98.6	1.4
17	Pd/C	$SnCl_2$	1	30	22	>99.9	>99.9	<0.1
18	Pd/C	$SnCl_2$	1	100	8	>99.9	98.9	1.1
19	Pd/Al_2O_3	–	1	30	12	15.5	89.5	10.5
20	Pd/Al_2O_3	$AlCl_3$	1	30	8	>99.9	>99.9	<0.1
21	Pd/Al_2O_3	$AlCl_3$	1	50	3	>99.9	99.3	0.7
22	Pd/Al_2O_3	$AlCl_3$	1	80	2	>99.9	99.0	1.0
23	Pd/NaY	–	1	30	12	10.6	82.3	17.7
24	Pd/NaY	$AlCl_3$	1	30	9	>99.9	>99.9	<0.1
25	Pd/NaY	$AlCl_3$	1	50	4	>99.9	>99.9	<0.1
26	Pd/NaY	$AlCl_3$	1	80	3	>99.9	99.3	0.7

*Second run to test the reusability of Pd/C-$ZnCl_2$ under the conditions of entry 16. †Third run to test the reusability of Pd/C-$ZnCl_2$ under the conditions of entry 16. ‡ Fourth run to test the reusability of Pd/C-$ZnCl_2$ under the conditions of entry 16.

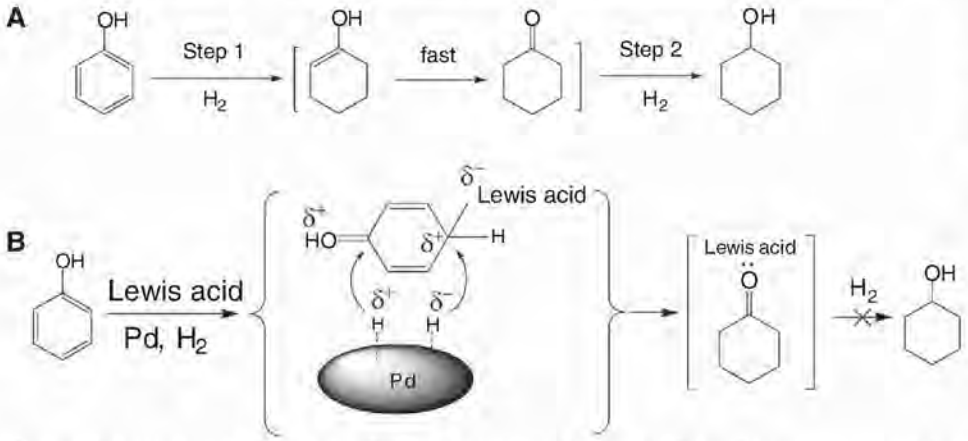


Fig 1. (**A**) General reaction pathway for hydrogenation of phenol. (**B**) Possible mechanism of dual activation in phenol hydrogenation and stabilization of cyclohexanone by Lewis acid.

It is known that Pd nanoparticles themselves are not active for the hydrogenation of benzene at low temperature (*25*), and they have very low activity for phenol hydrogenation (Table 1, entries 1, 19, 23). It has also been reported that Lewis acids can activate aromatic rings (*26*–*28*). Electrophilic aromatic substitution reactions have been extensively studied using experimental and theoretical techniques. Koltunov *et al.* (*26*) studied the reaction of 5-amino-1-naphthol with benzene in the presence of $AlCl_3$. Tarakeshwar *et al.* (*28*) studied the role of Lewis acids theoretically by performing high-level ab initio calculations on two model systems: C_6H_6-BCl_3 and C_6H_6-$AlCl_3$. Their results clearly indicated that one of the carbons in benzene tends to become highly nucleophilic, thereby facilitating attack of the benzene ring by an incipient electrophile. Furthermore, it is well known that Pd can activate H_2 (*29*). Based on the experimental data presented here and the literature results, we can conclude that in the hydrogenation of phenol, these two types of activation work cooperatively, resulting in high activity for producing cyclohexanone. This synergy also enhances the selectivity, because the reaction in the first step becomes relatively faster than that in the second step.

Inhibition of cyclohexanone hydrogenation (Fig. 1A, step 2) is another crucial factor underlying the high selectivity that we observed. We conducted the hydrogenation of cyclohexanone under some typical conditions using Pd/C and Pd/C-$AlCl_3$ as the catalysts (Table 3). The conversion of cyclohexanone catalyzed by Pd/C-$AlCl_3$ is much lower than that catalyzed by Pd/C. This result indicates that the Lewis acid suppresses the second step in Fig. 1A effectively. To account for this phenomenon, we studied the interaction between cyclohexanone and $AlCl_3$ in dichloromethane by Fourier transform infrared spectroscopy (fig. S7 and SOM text). The absorption band of the C=O stretching vibration shifts from 1714 cm^{-1} in the absence of $AlCl_3$ to 1624 cm^{-1} in the presence of $AlCl_3$. This shift is consistent with coordination of the Lewis basic C=O group to the Lewis acid (*30*). We believe that this Lewis acid–base interaction inhibits further hydrogenation of cyclohexanone.

Table 2. Hydrogenation of phenol in CO_2 under different pressures. Parameters were determined as in Table 1, and no additional products were observed. Reaction conditions were as follows: temperature, 30°C; p_{H_2}, 1.0 MPa; phenol, 2.0 mmol; Pd (5 wt % in Pd/C), 5 mol % relative to phenol; Lewis acid, 10 mol % relative to phenol.

Entry	Lewis acids	$P_{H_2} + P_{CO_2}$ (MPa)	Time (hours)	Conversion (%)	Selectivity (%)	
					Cyclohexanone	Cyclohexanol
1	$AlCl_3$	5.0	17.0	49.7	98.9	1.1
2	$AlCl_3$	6.5	13.0	63.9	98.2	1.8
3	$AlCl_3$	7.0	1.7	>99.9	>99.9	<0.1
4	$AlCl_3$	7.5	3.0	>99.9	>99.9	<0.1
5	$AlCl_3$	8.0	4.0	>99.9	>99.9	<0.1
6	$AlCl_3$	8.5	5.5	>99.9	>99.9	<0.1
7	$InCl_2$	7.0	8.0	>99.9	>99.9	<0.1
8	$InCl_2$	7.5	9.5	>99.9	>99.9	<0.1
9	$InCl_2$	8.0	10.0	>99.9	>99.9	<0.1
10	$ZnCl_2$	7.0	7.0	>99.9	>99.9	<0.1
11	$ZnCl_2$	7.5	7.3	>99.9	>99.9	<0.1
12	$ZnCl_2$	8.0	8.0	>99.9	>99.9	<0.1
13*	$AlCl_3$	8.0	4.0	>99.9	>99.9	<0.1
14†	$AlCl_3$	8.0	4.0	>99.9	>99.9	<0.1
15‡	$AlCl_3$	8.0	4.0	>99.9	>99.9	<0.1

*Second run to test the reusability of Pd/C-$AlCl_3$ under the conditions of entry 5. †Third run to test the reusability of Pd/C-$AlCl_3$ under the conditions of entry 5. ‡Fourth run to test the reusability of Pd/C-$AlCl_3$ under the conditions of entry 5.

Table 3. Hydrogenation of cyclohexanone under different conditions. Conversions and yields were determined by GC. Reaction conditions were as follows: cyclohexanone, 1.0 mmol; p_{H_2}, 1.0 MPa; Pd (5 wt % in Pd/C), 5 mol % relative to cyclohexanone; $AlCl_3$ (if used), 10 mol % relative to cyclohexanone; solvent, 1 ml of dichloromethane.

Entry	Catalysts	*T* (°C)	Time (hours)	Conversion (%)	Yield of cyclohexanol (%)
1	Pd/C	30	12	0.7	0.7
2	Pd/C-$AlCl_3$	30	12	<0.1	<0.1
3	Pd/C	50	7	12.9	12.9
4	Pd/C-$AlCl_3$	50	7	0.4	0.4
5	Pd/C	100	0.5	37.3	37.3
6	Pd/C-$AlCl_3$	100	0.5	1.8	1.8
7	Pd/C	100	4	53.7	53.7
8	Pd/C-$AlCl_3$	100	4	2.8	2.8

On the basis of the above discussion, the high activity and selectivity of the dual catalyst systems can be explained reasonably, as shown in Fig. 1B. The Pd, Lewis acid, and reactants contact each other when stirred. Lewis acid coordination makes the benzene ring of phenol more active, while Pd activates H_2, and cyclohexanone is formed quickly. At the same time, acid-base interaction between the Lewis acid and cyclohexanone inhibits further hydrogenation to cyclohexanol.

References and Notes

1. I. Dodgson, K. Griffen, G. Barberis, F. Pignataro, G. Tauszik, *Chem. Ind.* **1989**, 830 (1989).
2. M. Howe-Grant, Ed., *Kirk-Othmer Encyclopedia of Chemical Technology* (Wiley, New York, 1991).
3. L. X. Xu, C. H. He, M. Q. Zhu, K. J. Wu, Y. L. Lai, *Catal. Commun.* **9**, 816 (2008).
4. L. Liu *et al.*, *Angew. Chem. Int. Ed.* **48**, 2206 (2009).
5. S. G. Shore, E. Ding, C. Park, M. A. Keane, *Catal. Commun.* **3**, 77 (2002).
6. L. M. Sikhwivhilu, N. J. Coville, D. Naresh, K. V. R. Chary, V. Vishwanathan, *Appl. Catal. Gen.* **324**, 52 (2007).
7. S. Scirè, S. Minicò, C. Crisafulli, *Appl. Catal. Gen.* **235**, 21 (2002).
8. P. Claus *et al.*, *J. Catal.* **192**, 88 (2000).
9. E. Grasshoff *et al.*, Derwent Patent DD218092-A (1985).
10. C. V. Rode, U. D. Joshi, O. Sato, M. Shirai, *Chem. Commun. (Camb.)* **2003**, 1960 (2003).
11. H. J. Wang, F. Y. Zhao, S. Fujita, M. Arai, *Catal. Commun.* **9**, 362 (2008).
12. P. Makowski, R. D. Cakan, M. Antonietti, F. Goettmann, M. M. Titirici, *Chem. Commun. (Camb.)* **2008**, 999 (2008).
13. F. Lu, J. Liu, J. Xu, *Mater. Chem. Phys.* **108**, 369 (2008).
14. J. Morales, R. Hutcheson, C. Noradoun, I. F. Cheng, *Ind. Eng. Chem. Res.* **41**, 3071 (2002).
15. H. Li *et al.*, *Adv. Funct. Mater.* **18**, 3235 (2008).
16. E. Díaz *et al.*, *Chem. Eng. J.* **131**, 65 (2007).
17. C. Park, M. A. Keane, *J. Colloid Interface Sci.* **266**, 183 (2003).
18. Materials, methods, and further analysis are available as supporting material on *Science* Online.
19. M. Poliakoff, P. Licence, *Nature* **450**, 810 (2007).
20. P. G. Jessop, W. Leitner, *Chemical Synthesis Using Supercritical Fluids* (Wiley-VCH, Weinheim, Germany, 1999).
21. Z. F. Zhang *et al.*, *Angew. Chem. Int. Ed.* **47**, 1127 (2008).
22. M. G. Hitzler, M. Poliakoff, *Chem. Commun. (Camb.)* **1997**, 1667 (1997).
23. L. Zhuang, H. X. Li, W. L. Dai, M. H. Qiao, *Chem. Lett.* **32**, 1072 (2003).
24. S. Velu, M. P. Kapoor, S. Inagaki, K. Suzuki, *Appl. Catal. Gen.* **245**, 317 (2003).
25. C. Bianchini *et al.*, *Angew. Chem. Int. Ed.* **42**, 2636 (2003).
26. K. Y. Koltunov, G. K. S. Prakash, G. Rasul, G. A. Olah, *Tetrahedron* **58**, 5423 (2002).
27. R. R. Deshmukh, J. W. Lee, U. S. Shin, J. Y. Lee, C. E. Song, *Angew. Chem. Int. Ed.* **47**, 8615 (2008).
28. P. Tarakeshwar, J. Y. Lee, K. S. Kim, *J. Phys. Chem. A* **102**, 2253 (1998).
29. M. Benkhaled *et al.*, *Appl. Catal. Gen.* **346**, 36 (2008).
30. D. Cook, *Can. J. Chem.* **41**, 522 (1963).

31. We thank the National Natural Science Foundation of China (grants 20773144, 20733010, and 20633080), the National Key Basic Research Project of China (grant 2006CB202504), and the Chinese Academy of Sciences (grant KJCX2.YW.H16) for financial support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1250/DC1
Materials and Methods
SOM Text
Figs. S1 to S10
Tables S1 and S2
References

27 July 2009; accepted 30 September 2009
10.1126/science.1179713

Climate-Driven Basin-Scale Decadal Oscillations of Oceanic Phytoplankton

Elodie Martinez,* David Antoine, Fabrizio D'Ortenzio, Bernard Gentili

Phytoplankton—the microalgae that populate the upper lit layers of the ocean—fuel the oceanic food web and affect oceanic and atmospheric carbon dioxide levels through photosynthetic carbon fixation. Here, we show that multidecadal changes in global phytoplankton abundances are related to basin-scale oscillations of the physical ocean, specifically the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation. This relationship is revealed in ~20 years of satellite observations of chlorophyll and sea surface temperature. Interaction between the main pycnocline and the upper ocean seasonal mixed layer is one mechanism behind this correlation. Our findings provide a context for the interpretation of contemporary changes in global phytoplankton and should improve predictions of their future evolution with climate change.

Knowledge of decadal-scale changes in the distribution and abundance of ocean phytoplankton at a global scale remains uncertain. These changes result from processes occurring at a variety of time scales, including (possibly) the centennial one of anthropogenic global warming, interdecadal ones at basin-to-global scales, and interannual to seasonal ones. Understanding how processes intermingled at these different scales modify phytoplankton distributions requires global, multidecadal time series of chlorophyll concentration (Chl). This concentration describes at first order changes in phytoplankton biomass and also reflects the photoacclimation state of phytoplankton; that is, the modification of intracellular Chl in response to the average light level in the upper layers of the ocean. The best available tool for a global description of Chl is the remote sensing of ocean color from satellites, which allows multiyear global time series to be built from repeated synoptic observations.

Existing ocean color archives are relatively short in duration, however, so the long-term trend that may result from anthropogenic warming is neither unambiguously identified nor quantified. Contradictory results have been published as to whether the global ocean biomass increased or decreased from the 1980s to the 2000s (*1*, *2*). On the other hand, interannual-to-seasonal time scales have been intensively investigated on scales ranging from regional to global (*3*–*5*), in particular thanks to the 10-year time series provided by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) (*6*).

We focused on the multidecadal scale, by using a consistent, reanalyzed (*1*), ocean color time series built from 5 years of observations of the Coastal Zone Color Scanner (CZCS; 1979–1983) (*7*, *8*) and 5 years from SeaWiFS (1998–2002) (fig. S1). In parallel, a record of sea surface temperature (SST) (*8*, *9*) encompassing the same time period was used as an indicator of ocean stratification. These data sets allow us to perform a global quantification of the covariability of Chl and SST over two decades. We first describe how the inverse relationship between Chl and SST variabilities previously identified at interannual time scales for the SeaWiFS era (*4*) is also found over the decadal scale, but that the spatial distribution in the sign of these changes at the decadal scale is approximately a mirror image of those found over the SeaWiFS record. Multivariate empirical orthogonal function analyses (MEOFs) (*8*) were then performed in order to investigate the spatial and temporal signals behind this relationship, highlighting the role of basin-scale climate oscillators. Finally, other parameters were examined in order to propose a mechanism by which these oscillations of the physical environment affect the phytoplankton distribution.

Decadal changes in Chl and SST from the 1980s to the 2000s are displayed in Fig. 1, A and B, respectively. These maps highlight broad areas where Chl and SST evolved inversely. For the eastern and equatorial regions of the Pacific, SST decreased on average by 0.3°C, whereas in the western subtropical and western equatorial region, SST increased on average by 0.5°C (Fig. 1B). Changes in Chl, although more patchy than the SST changes, are roughly in the opposite direction, with large increases in the equatorial region where SST decreased. SST increased moderately (~0.2°C) in the Atlantic Ocean north of 30°S, and areas of inverse Chl and SST changes are less extensive there than in the Pacific. In the Indian Ocean, the areas where Chl and SST evolved inversely are mostly shaped by the distribution of the SST change, because Chl essentially increased everywhere in this ocean. Parallel increases of Chl and SST are the most frequent case (61%) in the Indian Ocean. This weak Chl-SST inverse relationship was also reported over the SeaWiFS era (*4*).

A map of the areas where Chl and SST evolve either in parallel (dark blue and purple) or inversely (light blue and yellow) is displayed in Fig. 1C. Overall, the regions of opposite Chl and SST changes represent about 60% of the ocean area comprised between latitudes 50°S and 50°N. This Chl-SST inverse correspondence was previously identified by (*4*) between 1999 and 2004 (Fig. 1D, similar to their figure 3C). The spatial distribution of the sign of these changes, as obtained here over two decades, is, however, approximately a mirror image of those found by (*4*) on interannual time scales. This appears in the Pacific when comparing Fig. 1, C and D. This inverted time evolution also appears regionally in the North Atlantic and barely shows up in the Indian Ocean. Therefore, when data over 1979–2002 (Fig. 1, A to C) or 1999–2004 (Fig. 1D) are analyzed, opposite directions of changes for SST and Chl are revealed in many regions. A large-scale multidecadal signal clearly emerges from this analysis, which indicates that oceanographic phenomena acting at basin scale shape the global decadal-scale phytoplankton distribution in the ocean.

The spatial patterns of the SST changes in the Pacific Ocean (Fig. 1B) actually bear the footprint of the Pacific Decadal Oscillation (PDO) (*10*), which is a pattern of SST anomalies in the subtropical North and South Pacific that are out of phase with SST anomalies in the tropical Pacific. In the Atlantic (Fig. 1B), a similar correspondence is observed with the Atlantic Multidecadal Oscillation (AMO) (*11*), whose pattern reveals SST anomalies that are homogeneous over the whole North Atlantic, positive for the period considered here. Other indexes such as the North Atlantic Oscillation or the Multivariate ENSO Index do not exhibit such spatial patterns, and their typical scales of variations are <~10 years, which makes them inappropriate for the present multidecadal analysis.

MEOFs were carried out over the 1979–2002 time period separately for the three oceanic basins (*8*). The reason for not doing a global analysis is that the Pacific represents ~50% of the global ocean and has the strongest SST and Chl signals, which overwhelm a global analysis, effectively masking the specific responses of the other oceans (*8*).

The spatial patterns of Chl and SST changes revealed in Fig. 1, A and B, also appear in the Chl and SST spatial variability maps made from the most significant components of the multidecadal signal (fig. S2). The Chl-SST common time variability associated with these spatial patterns is displayed for the three oceans in Fig. 2. In the Pacific Ocean (Fig. 2A), the signal shifts from negative to positive from 1979–1983 to 1998–2002, in correspondence with a shift from a PDO warm phase at the beginning of the 1980s to a cool phase at the end of the 1990s (*12*). Correlation between the PC and the PDO is 0.66 and reaches 0.83 when a 3-month lag between both is applied. There was a shift back to a PDO warm phase in 2002, which also appears in the Chl-SST time signal otherwise obtained over the 1998–2005 time period (right side of Fig. 2A; for a similar MEOFs analysis using the standard SeaWiFS

UPMC University of Paris 06, UMR 7093, Laboratoire d'Océanographie de Villefranche (LOV), 06230 Villefranche-sur-Mer, France; and CNRS, UMR 7093, LOV, 06230 Villefranche-sur-Mer, France.

*To whom correspondence should be addressed. E-mail: martinez@obs-vlfr.fr

Chl product, see fig. S3). This succession of opposite phases of the decadal oscillation corresponds to the opposite patterns obtained for the 1979–2002 (Fig. 1C) and 1999–2004 (Fig. 1D) analyses, for areas presenting opposite Chl and SST trends in this ocean (see also figs. S2 and S3). In the Indian Ocean, the Chl-SST covarying time signal also shifts from negative to positive from the beginning of the 1980s to the beginning of the 2000s (Fig. 2B). This shift may be related to the PDO [correlation coefficient (r) = 0.6] through the connection between the Indian and the Pacific. In the Atlantic, the sign of the time signal reverses from negative to positive from 1979–1983 to 1998–2002 (Fig. 2C). This corresponds to a shift of the AMO index from a cool to a warm phase in the 1990s (r = 0.75). On the contrary, the AMO index remains on average positive over the 1998–2005 period (right side of Fig. 2C). There is no shift from one phase of the decadal cycle to another, which explains why the spatial patterns in the Atlantic differ between Fig. 1C (1979–2002) and Fig. 1D (1999–2004) without being opposite as they are in the Pacific, however.

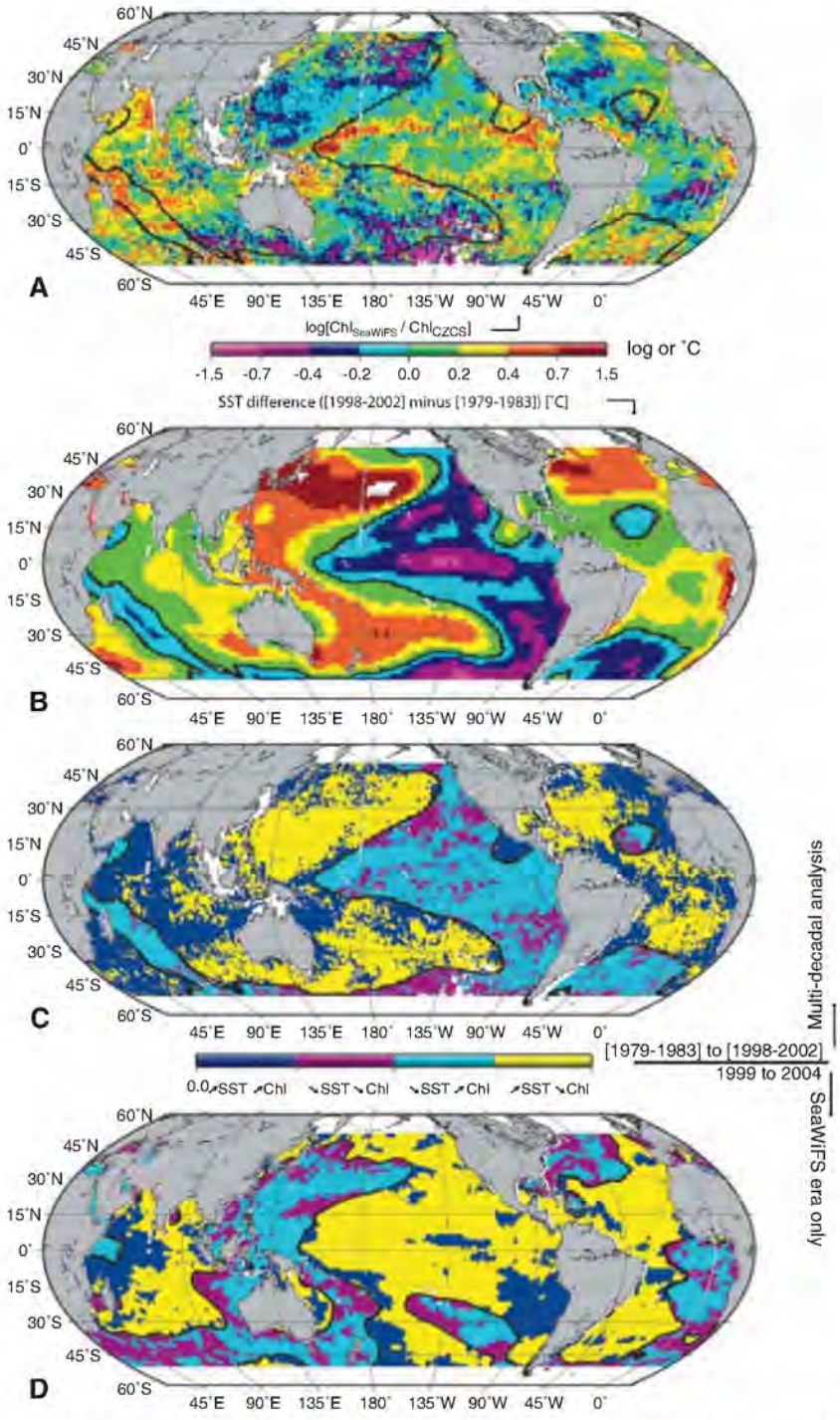


Fig. 1. (**A**) Chl change from the CZCS (1979–1983) to the SeaWiFS (1998–2002) era, expressed as the logarithm of the ratio of the average values over the two time periods [changes by a factor of 2 correspond to values of –0.7 (halving) or 0.7 (doubling)]. (**B**) SST difference over the same period. (**C**) Map of areas with concomitant parallel or opposite changes of Chl and SST, as indicated. (**D**) As in (C), but between 1999 and 2004, as in (*4*). The SST zero difference is shown on the maps as a thick black curve.

Correlations between basin-scale phytoplankton distribution and modes of climate oscillators are revealed here at multidecadal scale. In the tropics and at mid-latitudes, phytoplankton are often nutrient-limited. They grow when nutrients are made available within the upper lit layers through the upwelling of cold nutrient-rich water from below (or mixing with these deep waters). Therefore, any change of the overall stratification of the surface ocean has potential effects on phytoplankton growth, which may happen at different spatial and temporal scales. On seasonal-to-interannual time scales, there are changes of the depth of the upper ocean seasonal mixed layer (ZML). On decadal scales, the depth of the main pycnocline indicates the distance between the ZML (or the euphotic zone) and deep nutrient reservoirs. A proxy of the pycnocline is the depth of the 26 kg/m^3 isopycnal ($Z\rho_{26}$) (*13*), which is located between about 0 and 500 m within the intertropical area. Both ZML and $Z\rho_{26}$ were determined from the simple ocean data assimilation (SODA) model outputs (*8*, *14*).

The relationships between Chl and $Z\rho_{26}$ or ZML are displayed in Fig. 3 for the multidecadal analysis, similarly to the Chl-SST relationship shown in Fig. 1C. The patterns of $Z\rho_{26}$-Chl (Fig. 3A) and SST-Chl (Fig. 1C) show similarities, in particular in the Pacific and North Atlantic, showing that decadal-scale uplift or deepening of the pycnocline is the likely mechanism leading to the Chl-SST covariability. This inverse relationship is also found over almost the entire Indian Ocean. Changes in $Z\rho_{26}$ modify the average nutrient input to the upper lit layers, changing productivity and thereby chlorophyll levels. The geographical distribution of the Chl-ZML relationship (Fig. 3B) is not similar to the Chl-SST relationship shown in Fig. 1C. For instance, half of the area in the North Pacific where SST and Chl inversely evolve (yellow in Fig. 1C) shows parallel decreases in Chl and ZML (purple in Fig. 3B). The large areas of inverse Chl-SST in the equatorial and south Pacific are nearly absent from the Chl-ZML relationship. This shows that the Chl-ZML relationship probably bears the imprints of more-regional changes, intervening at shorter, interannual, scales.

Several studies have shown changes in global ocean phytoplankton over the past 10 years (*3*, *4*, *15*, *16*), and interpreting these changes as being indicative of long-term trends is tempting. Our results show that the multidecadal Chl changes observed over the CZCS-to-SeaWiFS era appear related to basin-scale oscillations in the physical environment (the PDO and AMO). These oscillations can alternately weaken or emphasize the possible effects of global warming (*17*), making difficult the identification of trends over short time series. In the Pacific, the PDO

results from the superposition of SST fluctuations forced in equal parts, at decadal time scales, by the variability of the Aleutian low, the El Niño–Southern Oscillation, and zonal advection anomalies in the Kuroshio-Oyashio Extension (*18*). In the Atlantic Ocean, the AMO has been linked to multidecadal variations in the strength of the meridional overturning circulation (*19*). Therefore, these processes all intervene to various degrees in shaping the global decadal-scale Chl variability, through their role on the overall stratification of the ocean, as revealed here through the

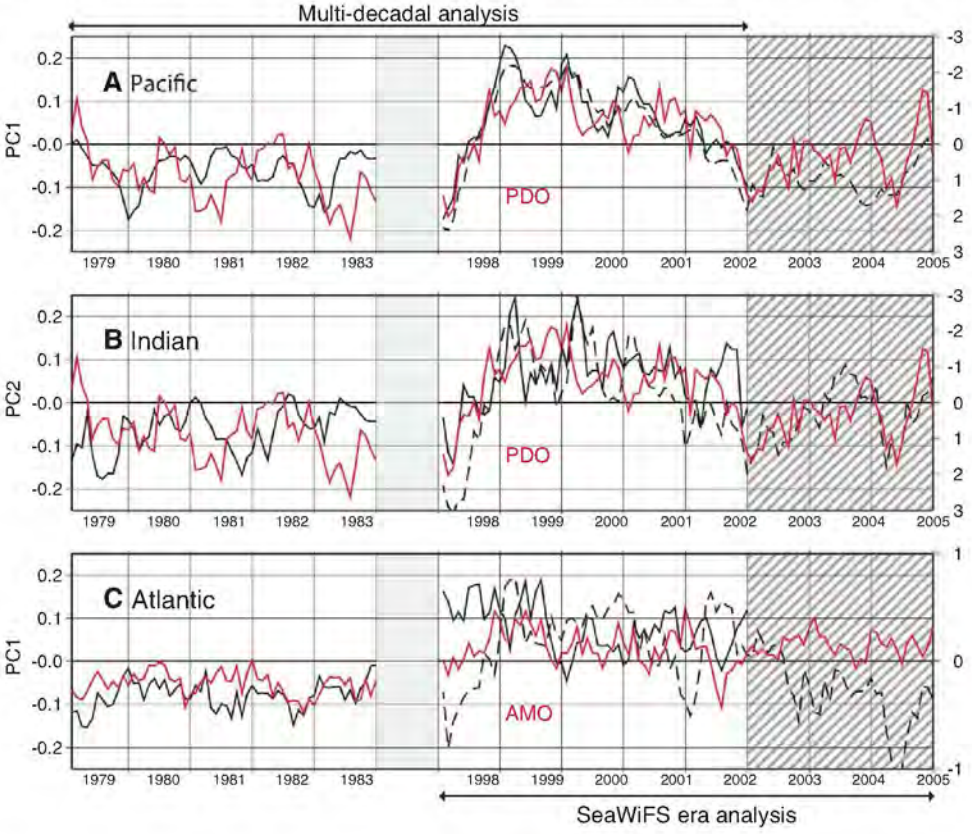


Fig. 2. The Chl-SST common time variability corresponding to the MEOF patterns (fig. S2) for the Pacific (**A**), Indian (**B**), and Atlantic (**C**) Oceans (black thick curve and scale on the left axes). PDO and AMO are superimposed as indicated (red curves and scales on the right axes). The Chl-SST common time variability issued from the MEOFs performed over the 1998–2005 era on the standard Chl SeaWiFS product (*8*) is shown as a dashed line on the right side of each panel. The diagonally hatched area sets the limit between the decadal analysis (using reanalyzed SeaWiFS data up to 2002 only) and the analysis presented in the supporting online material over the SeaWiFS era until 2005 (fig. S3).

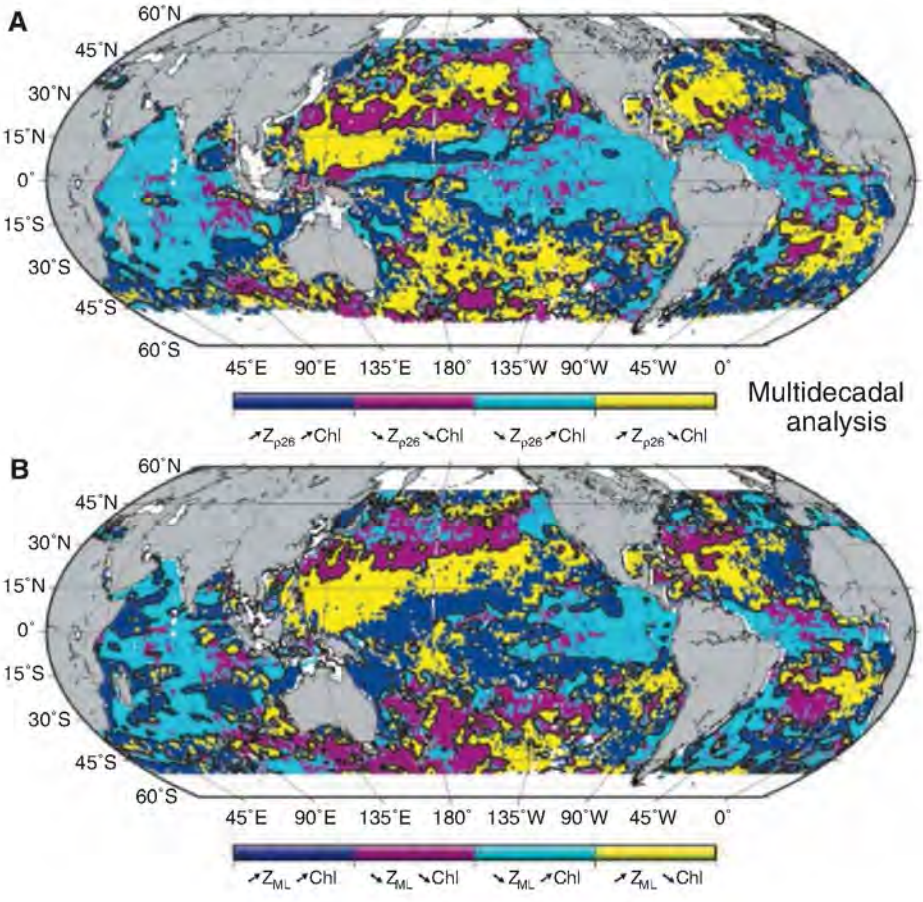


Fig. 3. Mapping of areas with concomitant parallel or opposite changes from 1979–1983 to 1998–2002 of Chl and $Z\rho_{26}$ (**A**), and of Chl and ZML (**B**), with $Z\rho_{\rho 26}$ and ZML determined from the SODA (*8*) data set. Dark blue and yellow correspond to deeper $Z\rho_{26}$ or ZML, whereas purple and light blue show areas of shallower $Z\rho_{26}$ or ZML.

Chl-$Z\rho_{26}$ and Chl-ZML relationships. This important role of ocean physics indicates that the observed Chl changes primarily reflect biomass changes due to dampened or increased nutrient fluxes to the upper lit layers. Changes in the photoacclimation state of phytoplankton probably also intervene, without obscuring the global picture, however. Quantifying the respective role of both phenomena would require the parallel examination of Chl and other quantities more directly tied to biomass.

A basin-specific response of phytoplankton to large-scale climate oscillators has been shown here. This result argues for a more accurate representation of decadal regimes into global ocean models, whose predictions of the response of ecosystems to global change are still uncertain (*20*–*22*). Such improvements are crucial for a better forecast of the impact of climate change on ecosystems and carbon fluxes. Our results also show that dampening the effect of interannual variability by averaging over two decades allows the decadal variability to be revealed and analyzed. Therefore, it can be anticipated that averaging over several decades may eventually reveal longer-term trends related to subtle changes in physical forcing. This emphasizes the critical importance of reanalyzing historical data sets (*23*, *24*) and of continuing the construction of climate-quality satellite data records in the next decades (*25*).

References and Notes

1. D. Antoine, A. Morel, H. R. Gordon, V. F. Banzon, R. H. Evans, *J. Geophys. Res.* **110**, C06009 (2005).
2. W. W. Gregg, M. E. Conkright, *Geophys. Res. Lett.* **29**, 1730 (2002).
3. C. R. McClain, S. R. Signorini, J. R. Christian, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **51**, 281 (2004).
4. M. J. Behrenfeld *et al.*, *Nature* **444**, 752 (2006).
5. J. A. Yoder, M. A. Kennelly, *Oceanography (Wash. D.C.)* **19**, 152 (2006).
6. C. R. McClain, G. C. Feldman, S. B. Hooker, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **51**, 5 (2004).
7. W. A. Hovis *et al.*, *Science* **210**, 60 (1980).
8. Materials and methods are available as supporting material on *Science* Online.
9. T. M. Smith, R. W. Reynolds, T. C. Peterson, J. Lawrimore, *J. Clim.* **21**, 2283 (2008).
10. N. J. Mantua, S. R. Hare, Y. Zhang, J. M. Wallace, R. C. Francis, *Bull. Am. Meteorol. Soc.* **78**, 1069 (1997).
11. D. Enfield, A. Mestas-Nunez, P. Trimble, *Geophys. Res. Lett.* **28**, 2077 (2001).
12. N. J. Mantua, S. R. Hare, *J. Oceanogr.* **58**, 35 (2002).
13. J. Hua, W. Dexing, W. Xiuquan, *Chin. J. Oceanol. Limnol.* **24**, 111 (2006).
14. J. A. Carton, B. S. Giese, *Mon. Weather Rev.* **136**, 2999 (2008).
15. S. C. Doney, *Nature* **444**, 695 (2006).
16. J. J. Polovina, E. A. Howell, M. Abecassis, *Geophys. Res. Lett.* **35**, L03618 (2008).
17. D. M. Smith *et al.*, *Science* **317**, 796 (2007).
18. N. Schneider, B. D. Cornuelle, *J. Clim.* **18**, 4355 (2005).
19. J. R. Knight, R. J. Allan, C. K. Folland, M. Vellinga, M. E. Mann, *Geophys. Res. Lett.* **32**, L20708 (2005).
20. L. Bopp *et al.*, *Global Biogeochem. Cycles* **15**, 81 (2001).
21. J. L. Sarmiento *et al.*, *Global Biogeochem. Cycles* **18**, GB3003 (2004).
22. B. Schneider, *Biogeosciences* **5**, 597 (2008).
23. R. H. Evans, H. R. Gordon, *J. Geophys. Res.* **99**, 7293 (1994).
24. P. G. Falkowski, C. Wilson, *Nature* **358**, 741 (1992).
25. C. R. McClain, S. B. Hooker, G. C. Feldman, P. Bontempi, *Eos* **87**, 337 (2006).

26. We thank the Agence Nationale de la Recherche (Paris) for the financial support for this work carried out within the frame of the Globphy project, NASA (U.S.) for providing the SeaWiFS global chlorophyll data, and the U.S. National Oceanographic and Atmospheric Administration National Climatic Data Center for the ERSST v3 data set. We thank D. Siegel, H. Claustre, Y. Huot, and A. Morel for the comments they provided on early versions of the manuscript, and three anonymous reviewers for their comments and helpful suggestions.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1253/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
References

29 May 2009; accepted 24 September 2009
10.1126/science.1177012

Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly

Michael E. Mann,[1]* Zhihua Zhang,[1] Scott Rutherford,[2] Raymond S. Bradley,[3] Malcolm K. Hughes,[4] Drew Shindell,[5] Caspar Ammann,[6] Greg Faluvegi,[5] Fenbiao Ni[4]

Global temperatures are known to have varied over the past 1500 years, but the spatial patterns have remained poorly defined. We used a global climate proxy network to reconstruct surface temperature patterns over this interval. The Medieval period is found to display warmth that matches or exceeds that of the past decade in some regions, but which falls well below recent levels globally. This period is marked by a tendency for La Niña–like conditions in the tropical Pacific. The coldest temperatures of the Little Ice Age are observed over the interval 1400 to 1700 C.E., with greatest cooling over the extratropical Northern Hemisphere continents. The patterns of temperature change imply dynamical responses of climate to natural radiative forcing changes involving El Niño and the North Atlantic Oscillation–Arctic Oscillation.

Considerable progress has been made over the past decade in using climate "proxy" data to reconstruct large-scale trends in past centuries, and in using climate models to assess the roles of natural and anthropogenic forcing in those trends (*1*). Owing in part to the sparseness of the available proxy data, less progress has been made in identifying the underlying spatial patterns of those changes, let alone the causal factors behind them. Yet a better understanding of past patterns of climate change and their causes (e.g., the role of past changes in the El Niño–Southern Oscillation, or ENSO) may be even more important for validating the regional-scale projections, which are paramount in assessing future climate change impacts.

Patterns of past climate change can be estimated through the simultaneous analysis of multiple spatially distributed proxy records. Such analyses have been performed via statistical reconstruction (*2*–*8*) and model assimilation approaches (*9*), but available proxy networks have been insufficient for estimating spatially resolved large-scale temperature reconstructions beyond the past few centuries (*2*, *4*, *7*).

Here, we employ a diverse multiproxy network previously used to estimate global and hemispheric mean annual temperature trends (*10*) to reconstruct global patterns of surface temperature changes over the past 1500 years. We use a climate field reconstruction (CFR) approach (*11*) that has been rigorously tested with synthetic "pseudoproxy" networks generated from forced climate model simulations (*12*). We interpret the resulting reconstructions in the context of results from climate model simulations forced by estimated past changes in natural (solar and volcanic) radiative forcing.

We employ the global proxy data set used by (*13*) comprising more than a thousand tree-ring, ice core, coral, sediment, and other assorted proxy records spanning the ocean and land regions of both hemispheres over the past 1500 years. The surface temperature field is reconstructed by calibrating the proxy network against the spatial information contained within the instrumental annual mean surface temperature field (*14*) over a modern period of overlap between proxy and instrumental data (1850 to 1995) using the RegEM CFR procedure (*12*) with additional minor modifications. Further details of

[1]Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA 16802, USA. [2]Department of Environmental Science, Roger Williams University, Bristol, RI 02809, USA. [3]Department of Geosciences, University of Massachusetts, Amherst, MA 01003–9298, USA. [4]Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ 85721, USA. [5]NASA Goddard Institute for Space Studies, New York, NY 10025, USA. [6]Climate Global Dynamics Division, National Center for Atmospheric Research, Boulder, CO 80305, USA.

*To whom correspondence should be addressed. E-mail: mann@meteo.psu.edu

the reconstruction procedure, associated statistical validation and skill assessments, uncertainty estimation procedures, data used, and MATLAB source codes for the analysis procedures are provided in the Materials and Methods.

Earlier proxy-based large-scale surface temperature reconstructions (*2*, *15*) resolved only a single statistical degree of freedom before the 15th century, precluding the possibility of investigating spatial patterns of surface temperature variation in earlier centuries. By contrast, the current reconstructions resolve multiple degrees of freedom in the surface temperature field back through the 6th century, allowing us to meaningfully interpret spatial features in the

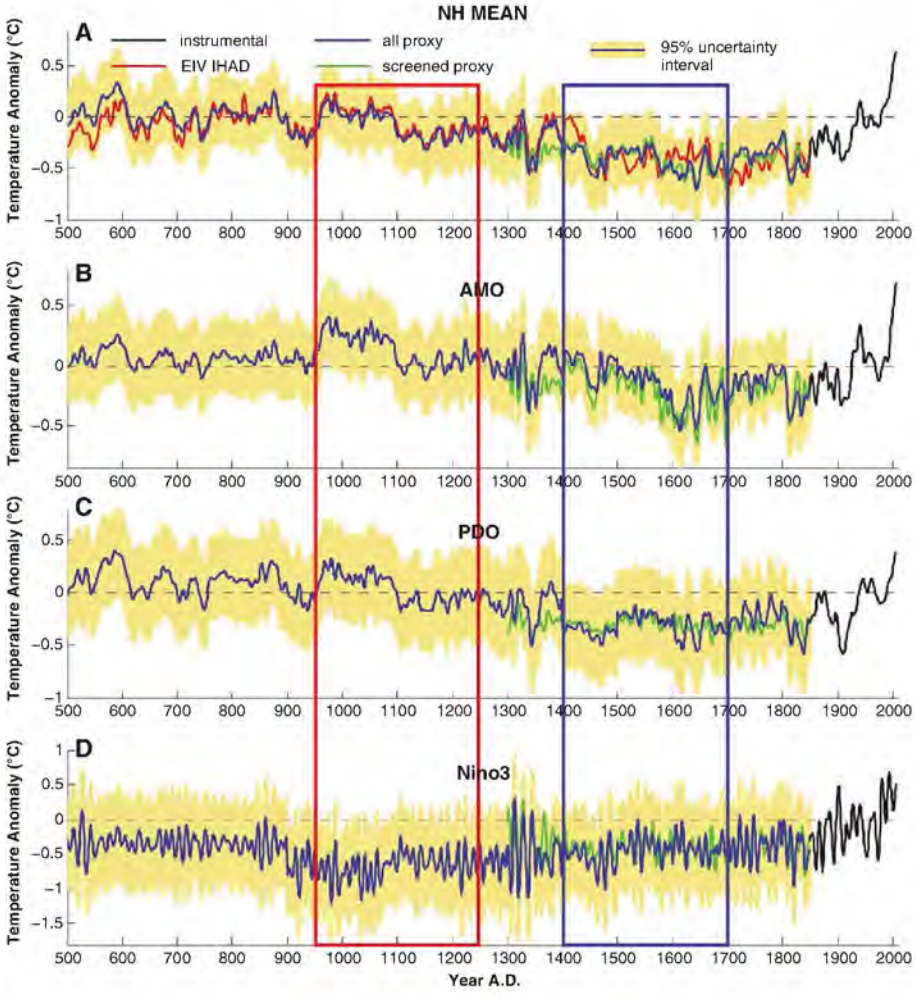


Fig. 1. Decadal surface temperature reconstructions. Surface temperature reconstructions have been averaged over (**A**) the entire Northern Hemisphere (NH), (**B**) North Atlantic AMO region [sea surface temperature (SST) averaged over the North Atlantic ocean as defined by (*30*)], (**C**) North Pacific PDO (Pacific Decadal Oscillation) region (SST averaged over the central North Pacific region 22.5°N–57.5°N, 152.5°E–132.5°W as defined by (*31*)], and (**D**) Niño3 region (2.5°S–2.5°N, 92.5°W–147.5°W). Shading indicates 95% confidence intervals, based on uncertainty estimates discussed in the text. The intervals best defining the MCA and LIA based on the NH hemispheric mean series are shown by red and blue boxes, respectively. For comparison, results are also shown for parallel ("screened") reconstructions that are based on a subset of the proxy data that pass screening for a local temperature signal [see (*13*) for details]. The Northern Hemisphere mean Errors in Variables (EIV) reconstruction (*13*) is also shown for comparison.

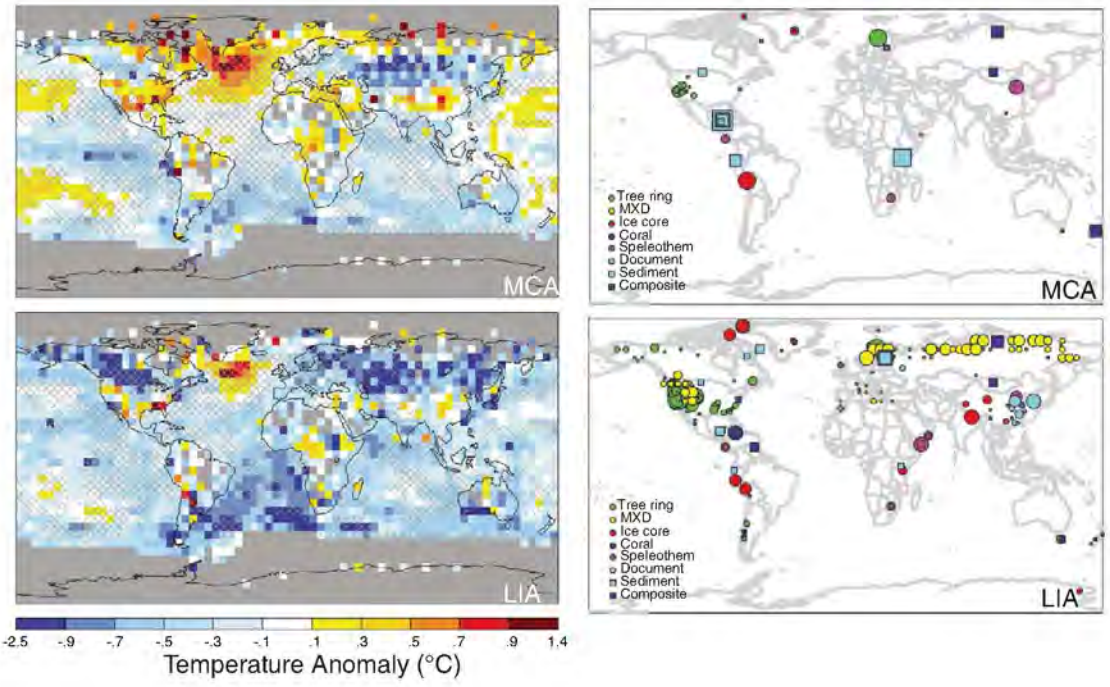


Fig. 2. Reconstructed surface temperature pattern for MCA (950 to 1250 C.E.) and LIA (1400 to 1700 C.E.). Shown are the mean surface temperature anomaly (left) and associated relative weightings of various proxy records used (indicated by size of symbols) for the low-frequency component of the reconstruction (right). Anomalies are defined relative to the 1961–1990 reference period mean. Statistical skill is indicated by hatching [regions that pass validation tests at the $P = 0.05$ level with respect to *RE* (*CE*) are denoted by / (\) hatching]. Gray mask indicates regions for which inadequate long-term modern observational surface temperature data are available for the purposes of calibration and validation.

reconstructions. Nonetheless, certain caveats must be kept in mind in interpreting the proxy-based surface temperature reconstructions. Before 1600 C.E., the low-frequency component of the surface temperature reconstructions is described as a linear combination of just two leading patterns of temporal variation, so that regional features in the temperature field are represented by a spatiotemporally filtered approximation. Moreover, as decadal-resolution proxy data were used in addition to annual-resolution data, only interdecadal and longer-term variations are meaningfully resolved; i.e., the details of individual years and even individual decades should not be emphasized. Thus, it is the longer-term, and larger-scale, variations resolved by the reconstructions that are most meaningful.

The large-scale surface temperature reconstructions, when spatially averaged, e.g., over the Northern Hemisphere, yield a long-term history very similar to the hemispheric mean reconstructions of (*13*) (Fig. 1A). However, the spatial reconstructions can also be averaged to yield other indices of interest [Fig. 1, B to E; other regional average series are shown in the Supporting Online Material (SOM) Text]. Though there are relatively few distinct patterns of variation resolved by the reconstructions, particularly before 1600 C.E., there are notable differences of behavior among the various diagnosed indices. The Atlantic Multidecadal Oscillation (AMO) series, for example, is marked by substantial multidecadal variability, consistent with previous proxy studies of North Atlantic variability [e.g., (*16*)]. The high-frequency fluctuations of the Niño3 series are consistent with the oscillatory nature of ENSO. The Niño3 index suggests strong and persistent La Niña conditions around 1000 years ago, as discussed further below.

Our reconstructions span two climatologically interesting periods, the so-called Little Ice Age (LIA) and Medieval Climate Anomaly (MCA). For the purpose of investigating the associated spatial patterns (Fig. 2), we defined the LIA and MCA in terms of distinct three-century-long intervals (1400 to 1700 C.E. and 950 to 1250 C.E., respectively), which both correspond to relative cold and warm hemispheric conditions, respectively (Fig. 1), and are distinct with regard to the estimated external radiative forcing of the climate (*1*, *17*). The observed patterns are not, however, sensitive to the precise time intervals used to define these periods (fig. S9). The MCA pattern is based on a smaller number of predictors than the LIA pattern (Fig. 2) and, accordingly, on fewer resolved spatial degrees of freedom (SOM Text).

The reconstruction skill diagnostics suggest that the MCA and LIA reconstructions are most reliable (Fig. 2) over the Northern Hemisphere and tropics, and least reliable in the Southern Hemisphere, particularly in the extratropics. To assess if the larger-scale features of the earlier MCA pattern are robust, we used only the more restricted network of proxy data available back through the beginning of the MCA interval to reconstruct temperatures for the LIA interval. This analysis gave a reconstruction very similar to the LIA reconstruction based on the full data set (fig. S10).

The reconstructed MCA pattern is characterized by warmth over a large part of the North Atlantic, Southern Greenland, the Eurasian Arctic, and parts of North America, which appears to

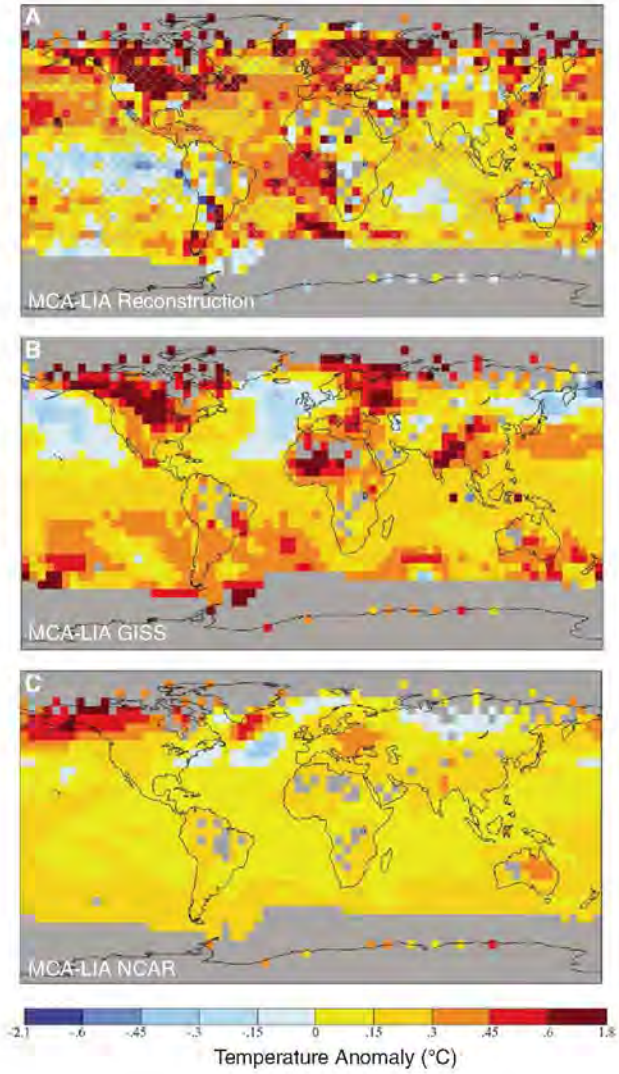


Fig. 3. Spatial pattern of MCA-LIA surface temperature difference in reconstructions and model simulations. (**A**) Proxy-based temperature reconstructions, (**B**) GISS-ER (using the same solar forcing difference used in the NCAR simulation—shown is the ensemble mean; see the SOM for example results from one of six realizations), and (**C**) NCAR CSM 1.4 simulation (using the same MCA and LIA time intervals as defined above). The observational mask has been applied to both model patterns for ease of comparison. Statistical skill for (A) is indicated with the same conventions as in Fig. 2 (statistical significance here indicates that the particular test statistic independently passed during both the MCA and LIA intervals).

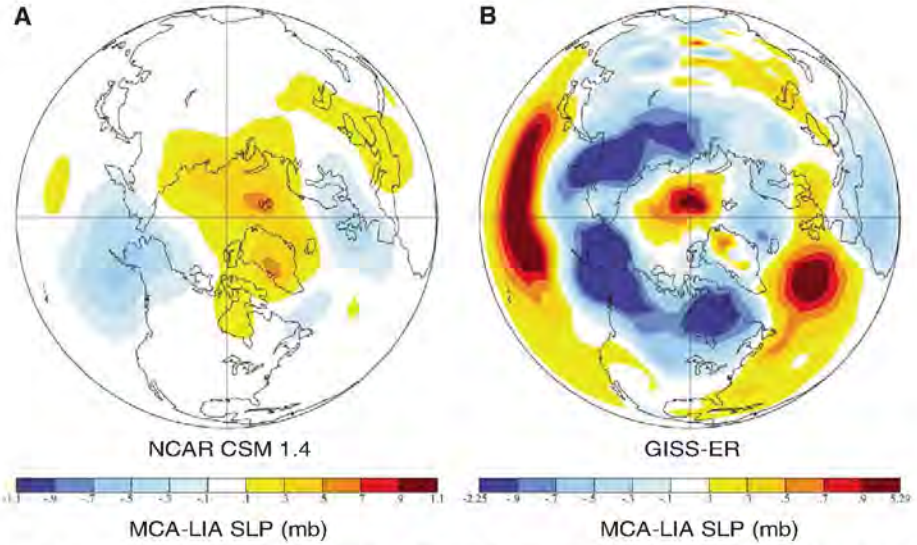


Fig. 4. Spatial pattern of MCA-LIA sea-level pressure difference in model simulations. (**A**) NCAR CSM 1.4 and (**B**) GISS-ER. For the NCAR model, a single run was available (*23*). For the GISS-ER coupled model, we show the ensemble mean of six realizations; see SOM section 5 for further details.

substantially exceed that of the modern late–20th century (1961–1990) baseline and is comparable to or exceeds that of the past one-to-two decades in some regions. This finding is consistent with that of a recent tree-ring–based study of high-latitude Eurasian temperatures (*18*). Relative warmth in the central North Pacific MCA is consistent with the expected extratropical signature of the strong observed La Niña–like pattern in the tropical Pacific (strong cooling in the east and warming in the west). Certain regions, such as central Eurasia, northwestern North America, and (with less confidence) parts of the South Atlantic, exhibit anomalous coolness. The LIA pattern is characterized primarily by pronounced cooling over the Northern Hemisphere continents, but with some regions—e.g., parts of the Middle East, central North Atlantic, Africa, and isolated parts of the United States, tropical Eurasia, and the extratropical Pacific Ocean—displaying warmth comparable to that of the present day. In some places, e.g., northern Labrador, apparent LIA warmth is a product, at least in part, of the relatively cool nature of the 1961–1990 reference period in the region.

For comparison with model simulation results, it is useful to eliminate the influence of the choice of modern reference period by examining the pattern of the MCA-LIA difference itself (Fig. 3). The MCA-LIA pattern highlights the extent to which the MCA is both more "La Niña–like" [e.g., (*17*, *19*–*21*)] and, with enhanced warmth over interior North America and the Eurasian Arctic, and cooling over central Eurasia, suggestive of the positive phase of the North Atlantic Oscillation (NAO) and closely related Arctic Oscillation (AO) sea-level pressure (SLP) pattern (*17*, *22*), as discussed further below.

We examined results for two different coupled model simulations of the past millennium, driven with those factors (solar irradiance changes and stratospheric aerosols from explosive volcanic eruptions) that can most plausibly explain the climate changes of the past millennium (*17*): (i) the National Center for Atmospheric Research (NCAR) Climate System Model (CSM) 1.4 coupled model driven with estimated solar plus volcanic forcing over the past millennium [see (*23*) for details]; and (ii) the Goddard Institute for Space Studies–ER (GISS-ER) coupled model with solar (but no volcanic) forcing (SOM Text), scaled for an MCA-LIA solar radiative forcing at the tropopause of 0.37 W/m^2 (equivalent to the MCA-LIA solar forcing difference used in the NCAR simulation). Both simulations give very similar estimates of the global mean MCA-LIA temperature difference (0.16° and 0.24°C for NCAR and GISS, respectively; the latter is identical to the proxy reconstructed mean surface temperature difference of 0.24°C). The spatial patterns of response for the two models (Fig. 3), however, are quite different, as discussed further below.

The La Niña–like nature of the MCA-LIA pattern is not reproduced in either of the two different coupled model simulations analyzed. On the other hand, such a pattern is reproduced in simulations (*19*) using the low-order Cane-Zebiak (*24*) model of the tropical Pacific coupled ocean-atmosphere system. The discrepancy in the model responses may arise because the tropical Pacific "thermostat" mechanism (*25*) is not active in either the NCAR or GISS simulations. In (*19*), this mechanism is responsible for the La Niña–like response to the positive tropical radiative forcing of the MCA that arises from a combination of relatively high solar irradiance and inactive tropical volcanism. Although there is still a vigorous debate regarding the nature of the response of the tropical Pacific to anthropogenic radiative forcing [e.g., (*26*)], paleoclimate evidence examined here, as elsewhere [e.g., (*19*, *27*)], appears to support a thermostat-like response, at least for natural radiatively forced climate changes in past centuries.

The NCAR simulation also does not reproduce the enhanced warming over the Eurasian Arctic, high-latitude North Atlantic, and North American region evident in the reconstructed MCA-LIA pattern. As discussed previously, this surface temperature pattern is consistent with a relative positive (negative) NAO-AO atmospheric circulation anomaly during the MCA (LIA), associated with annular bands of positive (negative) SLP anomalies in the subtropics and midlatitudes, and negative (positive) SLP anomalies in the subpolar latitudes. Such a pattern has been inferred in paleoclimate studies of the past millennium (*5*, *17*, *22*, *28*, *29*), and the negative phase of this pattern has been produced as a dynamical response to decreased solar radiative forcing during the LIA using a previous version of the NASA GISS model that incorporates the effects of ozone photochemistry on the vertical structure of the atmosphere (*28*, *29*). These effects are not accounted for in the NCAR simulation, which is limited to 36 km in vertical extent. The GISS-ER model used here extends to ~80 km and does incorporate these processes and, indeed, reproduces roughly the observed pattern of enhanced North American, high-latitude North Atlantic, and Arctic Eurasian warming, as a dynamical response to the imposed radiative forcing. These surface temperature changes are, in turn, associated with an annular atmospheric circulation response (Fig. 4) reminiscent of the positive phase of the NAO-AO pattern, though with some differences [in particular, (i) the high- and low-pressure regions in the North Atlantic sector are somewhat asymmetric and geographically shifted relative to the conventional pattern—hence, for example, the relative absence of warming in western Europe; and (ii) there is a positive SLP anomaly over Northern Greenland and part of the Eurasian Arctic Ocean that is absent in the conventional pattern]. Comparisons over the Pacific sector and neighboring regions, by contrast, are of limited utility, given the inability of the GISS-ER model to reproduce the aforementioned La Niña–like feature of MCA-LIA pattern, which strongly affects the Pacific basin. There is no evidence of a positive NAO-AO response in the NCAR simulation (Fig. 4).

The observed patterns of change, even when averaged over multicentury intervals, are unlikely to be entirely forced in nature, as there is also a potentially important role for purely internal, natural variability (*9*). Consistent with this view, we find that individual realizations of the GISS-ER transient response to the MCA-LIA solar forcing difference yield patterns that differ modestly in their details. For at least one realization, for example, the reconstructed warm anomaly over Western Europe is reproduced. In most cases, the basic features discussed above are nonetheless evident (SOM Text).

The paleoclimate reconstructions presented here hold important implications for future climate change. For example, if the tropical Pacific thermostat response suggested by our analyses of past changes applies to anthropogenic climate change, this holds profound implications for regional climate change effects such as future drought patterns. Continued refinement of paleoclimate reconstructions through expanded proxy databases and refinements of CFR methodology, improved estimates of past radiative forcing, and a better understanding of the influence of radiative forcing on large-scale climate dynamics should remain priorities as we work toward improving the regional credibility of climate model projections.

References and Notes

1. E. Jansen *et al.*, in *Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon *et al.*, Eds. (Cambridge Univ. Press, New York, 2007), pp. 433–497.
2. M. E. Mann, R. S. Bradley, M. K. Hughes, *Nature* **392**, 779 (1998).
3. K. R. Briffa, P. D. Jones, F. H. Schweingruber, T. J. Osborn, *Nature* **393**, 450 (1998).
4. K. R. Briffa *et al.*, *J. Geophys Res. Atmos.* **106**, 2929 (2001).
5. J. Luterbacher *et al.*, *Geophys. Res. Lett.* **26**, 2745 (1999).
6. J. Luterbacher, D. Dietrich, E. Xoplaki, M. Grosjean, H. Wanner, *Science* **303**, 1499 (2004).
7. M. N. Evans, A. Kaplan, M. A. Cane, *Paleoceanography* **17**, 1007 (2002).
8. S. Rutherford *et al.*, *J. Clim.* **18**, 2308 (2005).
9. H. Goosse, H. Renssen, A. Timmermann, R. S. Bradley, M. E. Mann, *Clim. Dyn.* **27**, 165 (2006).
10. See Dataset S1 in the Materials and Methods.
11. The synthetic proxy data in these tests are constructed to have noise characteristics similar to those estimated for actual proxy data. The calibration process, as with real-world reconstructions, is performed over a modern interval that is subject to anthropogenic forcing. The ability of the method to reproduce the earlier variations is then objectively assessed. See (*12*) for further details.
12. M. E. Mann, S. Rutherford, E. Wahl, C. Ammann, *J. Geophys. Res. Atmos.* **112**, D12109 (2007).
13. M. E. Mann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13252 (2008).
14. P. Brohan, J. J. Kennedy, I. Harris, S. F. B. Tett, P. D. Jones, *J. Geophys. Res. Atmos.* **111**, D12106 (2006).
15. M. E. Mann, R. S. Bradley, M. K. Hughes, *Geophys. Res. Lett.* **26**, 759 (1999).

16. T. L. Delworth, M. E. Mann, *Clim. Dyn.* **16**, 661 (2000).
17. P. D. Jones, M. E. Mann, *Rev. Geophys.* **42**, RG2002 (2004).
18. K. R. Briffa *et al.*, *Philos. Trans. R. Soc. B* **363**, 2271 (2008).
19. M. E. Mann, M. A. Cane, S. E. Zebiak, A. Clement, *J. Clim.* **18**, 447 (2005).
20. K. M. Cobb, C. D. Charles, R. L. Edwards, H. Cheng, M. Kastner, *Nature* **424**, 271 (2003).
21. N. E. Graham *et al.*, *Clim. Change* **83**, 241 (2007).
22. V. Trouet *et al.*, *Science* **324**, 78 (2009).
23. C. M. Ammann, F. Joos, D. Schimel, B. L. Otto-Bliesner, R. Tomas, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3713 (2007).
24. S. E. Zebiak, M. A. Cane, *Mon. Weather Rev.* **115**, 2262 (1987).
25. A. C. Clement, R. Seager, M. A. Cane, S. E. Zebiak, *J. Clim.* **9**, 2190 (1996).
26. G. A. Meehl *et al.*, in *Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon *et al.*, Eds. (Cambridge Univ. Press, New York, 2007), pp. 747–845.
27. J. Brad Adams, M. E. Mann, C. M. Ammann, *Nature* **426**, 274 (2003).
28. D. T. Shindell, G. A. Schmidt, M. E. Mann, D. Rind, A. Waple, *Science* **294**, 2149 (2001).
29. D. T. Shindell, G. A. Schmidt, R. L. Miller, M. E. Mann, *J. Clim.* **16**, 4094 (2003).
30. R. A. Kerr, *Science* **288**, 1984 (2000).
31. N. J. Mantua *et al.*, *Bull. Am. Meteorol. Soc.* **78**, 1069 (1997).
32. M.E.M. and Z.Z. gratefully acknowledge support from the ATM program of the National Science Foundation (grant ATM-0542356). R.S.B. acknowledges support from the Office of Science (BER), U.S. Department of Energy (grant DE-FG02-98ER62604). M.K.H. and F.B.N. were supported by the National Oceanic and Atmospheric Administration (grant NA16GP2914 from CCDD). D.T.S. and G.F. acknowledge support from NASA's Atmospheric Chemistry, Modeling, and Analysis Program.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1256/DC1
Materials and Methods
SOM Text
Figs. S1 to S11
Tables S1 to S5
References
SOM Data

4 June 2009; accepted 5 October 2009
10.1126/science.1177303

Extensive, Recent Intron Gains in *Daphnia* Populations

Wenli Li,[1]* Abraham E. Tucker,[1]* Way Sung,[2] W. Kelley Thomas,[2] Michael Lynch[1]†

Rates and mechanisms of intron gain and loss have traditionally been inferred from alignments of highly conserved genes sampled from phylogenetically distant taxa. We report a population-genomic approach that detected 24 discordant intron/exon boundaries between the whole-genome sequences of two *Daphnia pulex* isolates. Sequencing of presence/absence loci across a collection of *D. pulex* isolates and outgroup *Daphnia* species shows that most polymorphisms are a consequence of recent gains, with parallel gains often occurring at the same locations in independent allelic lineages. More than half of the recent gains are associated with short sequence repeats, suggesting an origin via repair of staggered double-strand breaks. By comparing the allele-frequency spectrum of intron-gain alleles with that for derived single-base substitutions, we also provide evidence that newly arisen introns are intrinsically deleterious and tend to accumulate in population-genetic settings where random genetic drift is a relatively strong force.

Introns are noncoding sequences that interrupt eukaryotic exons and are removed from premature mRNAs by the spliceosomal machinery before translation (*1*–*3*). Intron colonization affects the evolution of gene structure and is a factor in the emergence of genomic and organismal complexity, as newly arisen introns are thought to be intrinsically deleterious owing to the increased mutational target that they impose on their host genes (*4*, *5*). The number of introns in a genome is determined by the relative rates of intron gain and loss over evolutionary time, which differ among lineages. Across eukaryotes, intron numbers range from >100,000 per vertebrate genome to only two in *Giardia lamblia* (*6*, *7*). The fundamental causes of this variation remain controversial (*8*, *9*), partly because of a lack of population-level analyses with the power to infer the properties of recent gain or loss alleles.

The early eukaryotic progenitor has been assumed to be intron-rich on the basis of the presence of introns in homologous positions of orthologous genes of widely divergent eukaryotes (*10*–*12*) and the likely presence of a complex spliceosome in the eukaryotic ancestor (*13*). In this context, intron-poor lineages are assumed to reflect a long-term history of intron loss (*14*). Alternatively, moderate ancestral intron density followed by lineage-specific gains (*15*) may have occurred, even at orthologous positions in divergent taxa (*16*). However, most comparative studies of introns have examined only a small subset of highly conserved genes between deeply divergent lineages, and although some studies have documented unambiguous examples of intron gain (*17*–*19*) and some statistical procedures allow an indirect inference of parallel gains and/or losses (*20*, *21*), comparative studies of taxa with extreme sequence divergence have essentially no possibility of directly inferring parallel intron gains.

Because they potentially retain the molecular signatures of the process of intron origin, intron presence or absence alleles segregating in natural populations provide material to infer gain or loss mechanisms and to estimate taxon-specific turnover rates. Such polymorphisms do exist. A standing intron presence/absence polymorphism was found at a locus in natural isolates of *Drosophila teisseri* (*22*), and two intron-gain alleles segregate with an intron-free version at a locus in the microcrustacean *Daphnia pulex* (*23*). The latter study, in particular, inspired us to look more deeply for evidence of recent intron gain or loss in *D. pulex*.

By artificially removing intron sequences from all predicted gene sequences of the annotated *D. pulex* genome [clone TCO (*24*)] and querying the exon-exon boundaries (n = 110,021) against another *D. pulex* genome sequence (TRO), we detected putative intron-free alleles. After filtering for paralogy and false positives, such as processed pseudogenes, we sequenced the genomic regions surrounding 24 intron presence/absence positions across 84 natural isolates of North American *D. pulex* species as well as in eight *Daphnia* outgroup species. Gene trees constructed from flanking-exon sequence for each presence/absence polymorphism revealed the phylogenetic relationships of the polymorphic alleles, and from these data we inferred that 87.5% (21/24) of the intron polymorphisms reflect recent intron gains, with three reflecting intron losses (figs. S1 to S24). Most of the gains (15/28) were exclusive to Oregon populations, a genetically isolated subclade of North American *D. pulex* (*25*, *26*) with a historically low effective population size (*27*). Active splicing of all polymorphic introns was confirmed with reverse transcription polymerase chain reaction sequencing.

The features of newly arisen introns in *D. pulex* are inconsistent with most hypothesized mechanisms of intron origin (*7*). We found no support for intron gains resulting from tandem duplications of fragments of coding DNA or insertions of transposable elements. Furthermore, the polymorphic intron sequences identified seem to be evolutionary novelties absent from well-characterized eukaryotic and prokaryotic genomes. Except for gains at one locus (Dappu-42116_2, fig. S20), Blastn searches using recently gained intron sequences against the *D. pulex* genome assembly (http://wfleabase.org/blast), *D. pulex* genome trace files, and the full GenBank repository did not retrieve any homologous sequence hits.

We observed that short direct repeats, ranging in size from 5 to 12 base pairs, flank many

[1]Biology Department, Indiana University, Bloomington, IN 47405, USA. [2]Hubbard Center for Genome Studies, University of New Hampshire, Durham, NH 03824, USA.
*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: milynch@indiana.edu

Fig. 1. Sequence alignment showing three parallel intron gains in *D. pulex* at locus Dappu-42116_2. Homologous intron sequences are indicated by color-coded brackets; different gains are indicated by different colors; intron-containing and intron-absent alleles are indicated by plus and minus signs, respectively.

(12/28) of the intron gains, with one repeat positioned within the end of an adjacent exon and the other repeat near the opposite end of the intron sequence (figs. S1 to S3, S7 to S11, S15 to S18, and S21). These sequences suggest that intron gains in *D. pulex* result from recent repair of staggered double-strand breaks (DSBs) accompanied by small segmental insertions (either preexisting fragments or newly synthesized). In other systems, DSBs and subsequent nonhomologous end joining (NHEJ) repair are known to be associated with small insertions of exogenous DNA (*28*), including mitochondrial DNAs (*29*). Consistent with this model, one recent intron gain identified in our study was homologous to the 16*S* ribosomal subunit of the *D. pulex* mitochondrial genome, although, as noted above, the source of other gained introns remains unresolved. We also noted that the AT content of recently gained introns (80.9 ± 1.3%) was significantly higher than that of other, nonpolymorphic introns in the same genes (70.8 ± 0.5%), which are themselves high in AT relative to surrounding exons (54.1 ± 1.0%). This suggests that AT-rich insertion sequences are particularly prone to intronization.

In four of the loci, multiple intron origins have arisen at the exact same site (Fig. 1, Fig. 2, and figs. S15, S19, S20, and S21). If these observations were each due to a single intron-gain event followed by subsequent divergence of intronic sequence, we would expect an overall high rate of allelic divergence in the surrounding exons, but the latter exhibit typical levels of sequence divergence (no more than 3 to 6%), as do other introns in the same gene. Among the parallel intron gains identified, two involve a phase difference (figs. S15 and S19), where the intron interrupts a different position of the same codon. Thus, we conclude that the divergent intron alleles observed at homologous sites are independently derived. Because 4 of 21 total sites of intron gain harbor such

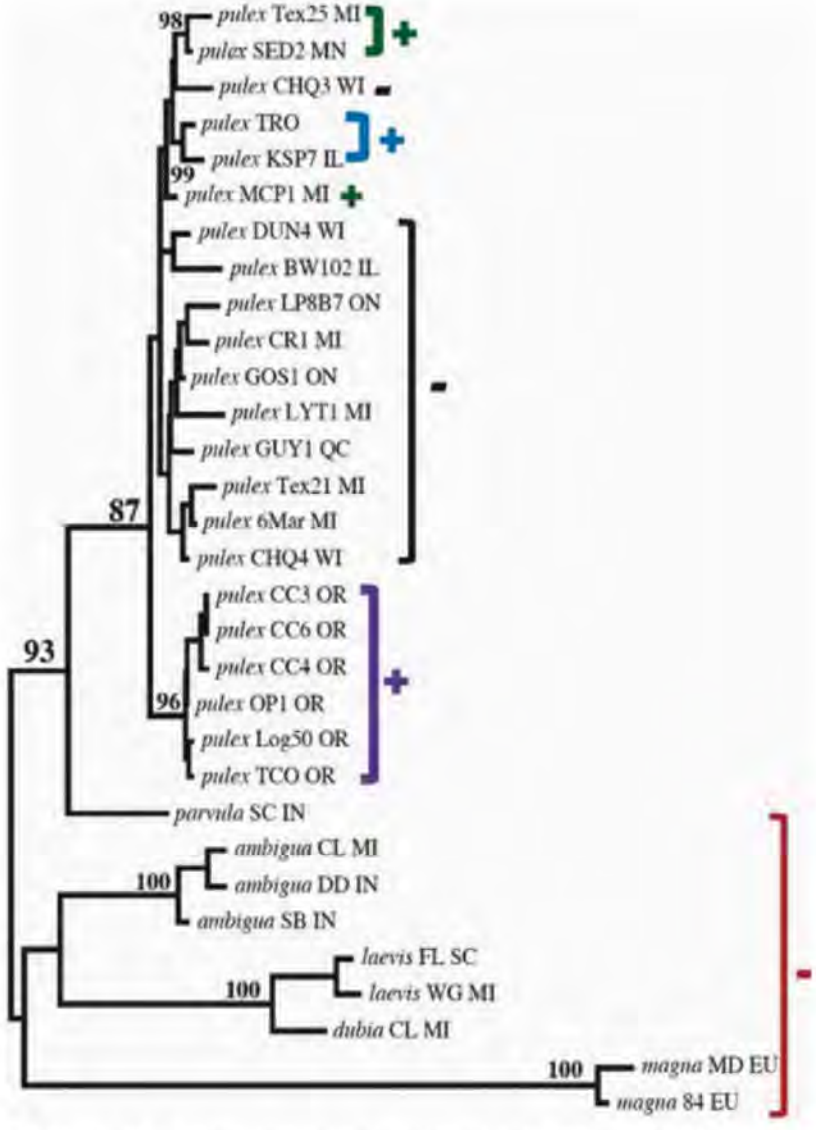


Fig. 2. Neighbor-joining gene tree with bootstrap values greater than 85% showing three parallel intron gains in *D. pulex* at locus Dappu-42116_2. Color codes and symbols are as in Fig. 1.

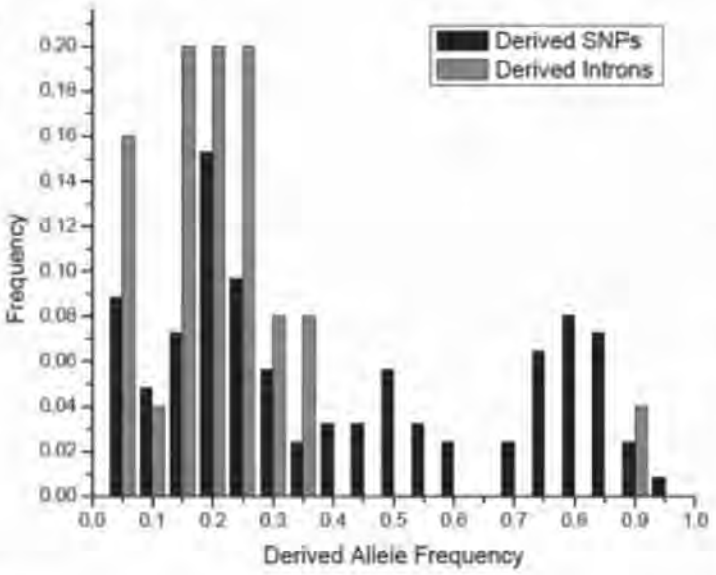


Fig. 3. Allele-frequency spectra of derived SNPs and introns in *D. pulex*. Gained introns are skewed to lower frequency than derived SNPs.

independent sequences, this further implies that intron-insertion hot spots may exist in the *D. pulex* genome.

A motif search of exon sequences immediately surrounding the polymorphic introns did not yield any motif common to the insertion sites. Whereas nearly half (45%) of established *D. pulex* introns reside in between-codon positions (phase 0), most of the gained introns in this study were inserted after the first positions within codons (phase 1) (*24*). Because introns that split codons are more prone to deleterious consequences during faulty splicing and/or intron sliding (*4*), they may be less likely to be fixed, which suggests that some percentage of intron gains are transient and not typical of established introns.

Purifying selection on derived introns relative to other base-substitution polymorphisms can be inferred by comparing the allele-frequency spectrum of derived single-nucleotide polymorphisms (SNPs) and gained introns (Fig. 3). We found that the allele-frequency distribution of recently gained introns is skewed to low levels relative to derived SNPs (average frequency of derived introns is 19.5 ± 2.6% versus 39.4 ± 3.4% for derived SNPs).

The presence of upstream and downstream introns argues against presence/absence polymorphisms being artifacts of processed pseudogenes, where adjacent introns are expected to be missing because of the incorporation of a fully spliced cDNA transcript. Moreover, this additional sequencing also yielded another intron presence/absence polymorphism within *D. pulex* that was not previously detected with our bioinformatic survey of TCO and TRO, which suggests that there may be many more allelic variants yet to be discovered. Such sequencing also uncovered intron gains in two outgroup species, *D. obtusa* and *D. parvula*, indicating that intron gains may not be exclusive to *D. pulex*. In addition, we detected a parallel intron gain at locus Dappu-323635_3 in *D. laevis* (fig. S16), consistent with the view that the existence of intron colonization hot spots is not restricted to *D. pulex*. Our study has an ascertainment bias, making identification of intron presence preferential in TCO-related populations, because the exon boundaries used to detect intron absences were generated from the predicted annotations of the TCO genome. However, any polarizations as a gain or loss are not affected by this bias.

Intron gain has been argued to be a rare event, with a rate on the order of $<4 \times 10^{-6}$ per coding site per million years, which is orders of magnitude lower than estimated rates of loss (*5*, *30*). Massive losses of ancestral introns have been postulated to have occurred in select lineages (*14*, *31*, *32*), and it has been suggested that rates of intron gain and loss have been declining in the past 1.3 billion years in most eukaryotes, with a greater decline of gains than losses (*32*). Regardless of whether these are accurate inferences, they are not consistent with our observations of *Daphnia* intron turnover, where the rate of gain is higher than the rate of loss and is minimally 1.2×10^{-5} per coding site per million years (*24*).

Our data suggest that rates of intron turnover, particularly intron gain, are higher than previously appreciated. Furthermore, the documentation of parallel gains occurring at many of the same sites is contrary to the assumptions in many prior analyses. If similar processes occur in other taxa, analyses of intron turnover rates that fail to account for rampant parallel gain may lead to underestimates of the rates of gain and overestimates of the rates of loss. In addition, our identification of the short direct repeats found in association with recently originated intron-containing alleles suggests that introns may be gained fortuitously as a consequence of DNA damage, with repair of staggered DSBs being occasionally accompanied by insertions that by chance harbor the sequences necessary to elicit a splicing reaction.

References and Notes

1. S. M. Berget, C. Moore, P. A. Sharp, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3171 (1977).
2. L. T. Chow *et al.*, *Cell* **12**, 1 (1977).
3. W. Gilbert, *Nature* **271**, 501 (1978).
4. M. Lynch, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6118 (2002).
5. M. Lynch, *The Origins of Genome Architecture* (Sinauer, Sunderland, MA, 2007).
6. J. E. J. Nixon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3701 (2002).
7. S. W. Roy, W. Gilbert, *Nat. Rev. Genet.* **7**, 211 (2006).
8. R. Belshaw, D. Bensasson, *Heredity* **96**, 208 (2006).
9. D. C. Jeffares, T. Mourier, D. Penny, *Trends Genet.* **22**, 16 (2006).
10. A. Fedorov, A. F. Merican, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16128 (2002).
11. I. B. Rogozin, Y. I. Wolf, A. V. Sorokin, B. G. Mirkin, E. V. Koonin, *Curr. Biol.* **13**, 1512 (2003).
12. J. E. Stajich, F. S. Dietrich, S. W. Roy, *Genome Biol.* **8**, R223 (2007).
13. L. Collins, D. Penny, *Mol. Biol. Evol.* **22**, 1053 (2005).
14. S. W. Roy, D. Penny, *Mol. Biol. Evol.* **24**, 1926 (2007).
15. S. W. Roy, D. Penny, *Mol. Biol. Evol.* **24**, 1447 (2007).
16. R. Tarrío, F. Rodríguez-Trelles, F. J. Ayala, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6580 (2003).
17. A. Coghlan, K. H. Wolfe, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11362 (2004).
18. T. Hankeln, H. Friedl, I. Ebersberger, J. Martin, E. R. Schmidt, *Gene* **205**, 151 (1997).
19. C. B. Nielsen *et al.*, *PLoS Biol.* **2**, e422 (2004).
20. L. Carmel, I. B. Rogozin, Y. I. Wolf, E. V. Koonin, *BMC Evol. Biol.* **7**, 192 (2007).
21. H. D. Nguyen, M. Yoshihama, N. Kenmochi, *PLOS Comput. Biol.* **1**, e79 (2005).
22. A. Llopart, J. M. Comeron, F. G. Brunet, D. Lachaise, M. Long, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8121 (2002).
23. A. R. Omilian, D. G. Scofield, M. Lynch, *Mol. Biol. Evol.* **25**, 2129 (2008).
24. See supporting material on *Science* Online.
25. J. K. Colbourne *et al.*, *Biol. J. Linn. Soc.* **65**, 347 (1998).
26. S. Paland, J. K. Colbourne, M. Lynch, *Evolution* **59**, 800 (2005).
27. M. Lynch *et al.*, *Evolution* **53**, 100 (1999).
28. X. Yu, A. Gabriel, *Mol. Cell* **4**, 873 (1999).
29. E. Hazkani-Covo, S. Covo, *PLoS Genet.* **4**, e1000237 (2008).
30. S. W. Roy, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5773 (2005).
31. S. Cho, S. Jin, A. Cohen, R. E. Ellis, *Genome Res.* **14**, 1207 (2004).
32. L. Carmel, Y. I. Wolf, I. B. Rogozin, E. V. Koonin, *Genome Res.* **17**, 1034 (2007).

33. Supported by NSF grants MCB-0342431, EF-082741, and EF-0328516 (M.L.). We thank F. Catania, T. Doak, J. Colbourne, and K. Montooth for discussions, and A. Seyfert, S. Schaack, A. Omilian, and E. Williams for early contributions to this project. Genome sequences and annotation are available through the collaboration of JGI and the Daphnia Genomics Consortium at http://genome.jgi-psf.org/Dappu1/Dappu1.home.html. Intron data are provided in the online supplement. Sequences have been deposited in GenBank, accession numbers GQ984366 to GQ985204 (for details see tables S6 and S7).

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1260/DC1
Materials and Methods
Figs. S1 to S25
Tables S1 to S7
References

17 July 2009; accepted 24 September 2009
10.1126/science.1179302

Impact of Genome Reduction on Bacterial Metabolism and Its Regulation

Eva Yus,[1] Tobias Maier,[1] Konstantinos Michalodimitrakis,[1] Vera van Noort,[2] Takuji Yamada,[2] Wei-Hua Chen,[2] Judith A. H. Wodke,[1] Marc Güell,[1] Sira Martínez,[1] Ronan Bourgeois,[1] Sebastian Kühner,[2] Emanuele Raineri,[1] Ivica Letunic,[2] Olga V. Kalinina,[2,3] Michaela Rode,[2] Richard Herrmann,[3] Ricardo Gutiérrez-Gallego,[4] Robert B. Russell,[2] Anne-Claude Gavin,[2] Peer Bork,[2]* Luis Serrano[1,6]

To understand basic principles of bacterial metabolism organization and regulation, but also the impact of genome size, we systematically studied one of the smallest bacteria, *Mycoplasma pneumoniae*. A manually curated metabolic network of 189 reactions catalyzed by 129 enzymes allowed the design of a defined, minimal medium with 19 essential nutrients. More than 1300 growth curves were recorded in the presence of various nutrient concentrations. Measurements of biomass indicators, metabolites, and ^{13}C-glucose experiments provided information on directionality, fluxes, and energetics; integration with transcription profiling enabled the global analysis of metabolic regulation. Compared with more complex bacteria, the *M. pneumoniae* metabolic network has a more linear topology and contains a higher fraction of multifunctional enzymes; general features such as metabolite concentrations, cellular energetics, adaptability, and global gene expression responses are similar, however.

Accurate representation of cellular networks through mathematical models is a central goal of integrative systems biology. For this purpose, all components and reactions of a target system should be listed and validated, and their quantitative relations should be determined and analyzed in the context of the physiology of the organism (*1*). We have selected *Mycoplasma pneumoniae*, a human pathogen that causes atypical pneumonia (*2*), as a model organism for bacterial and archaeal systems biology. Similar to other Mollicutes, *M. pneumoniae* has undergone a massive genome reduction to include only 689 protein coding genes, 231 of which have unknown function (table S1) (*3*), yet it can be cultivated in vitro without helper cells (*4*). The genome reduction of *M. pneumoniae* favors its suitability as a systems biology model because it largely follows genome size–scaling principles (fig. S1) (*5*). We manually reconstructed and validated the metabolic network of *M. pneumoniae* and studied its regulation, complementing analyses of the transcriptome (*6*) and the proteome organization (*7*).

The metabolism of *M. pneumoniae* has been studied biochemically (*8*) and computationally (*9*). We integrated these approaches in a framework that maximized coverage and accuracy (*10*). To build a comprehensive metabolic network, we complemented the reactions from the Kyoto Encyclopedia of Genes and Genomes (KEGG; www.genome.jp/kegg) with activities obtained manually from the literature and new annotations (fig. S2 and tables S1 to S5) (*11*). We also considered other genomic (co-occurrence in one operon), sequence (homology to known enzymes), and structural information (identification of catalytic residues to ensure enzyme functionality) (Fig. 1A and figs. S2 and S3). For example, we identified an incomplete ascorbate pathway through sequence analyses and filled the gap by assigning a critical enzyme [L-ascorbate-6-phosphate lactonase (*mpn497*)] on the basis of sequence homology, predicted activity (metal-dependent hydrolase), and its position in the ascorbate operon (*mpn492* to *mpn497*). For pathways in which only one enzyme was missing, we closed the gap by adding an unassigned reaction (for example, transketolase activity in the pentose phosphate pathway). Putative enzymes missing conserved catalytic residues were discarded (for example, Mpn255 and Mpn673 enzymes of the terpenoid pathway). Lastly, for enzymes that could carry out more than one reaction, we removed the reactions that were decoupled from pathways and those for which the substrate was unavailable. The final result was a map without gaps, isolated reactions, or open metabolic loops (Fig. 2).

A number of alternative pathways, interactions between pathways, as well as missing enzymes still needed to be validated, and reaction directionalities had to be inferred. For this, we used two different experimental strategies. We first used the rich medium (fig. S4) to validate the pathway functionality in various carbon sources. As expected from the map (Fig. 2), all known carbon sources except mannitol supported growth to various extents (figs. S5 and S6) (*12*). Using ^{13}C-glucose labeling, we validated (for example) the predicted connection between glycolysis, the pentose phosphate pathway, and lipid synthesis (fig. S3 and table S6), and ruled out the proposed production of aspartate from pyruvate (*13*). For our second strategy, we developed on the basis of the metabolic map a defined medium (Fig. 1A and table S7) with which we could validate other pathways (such as vitamin metabolism) (fig. S10) and reaction directionalities that could not be studied in rich medium (such as the synthesis of uracyl and thymine nucleotides from cytosine) (figs. S7 and S8). The low number of amino acid permeases and transporters and the existence of a peptide importer (*oppB-F* cluster) (table S1) suggested a requirement for peptides in the medium, which we confirmed experimentally (fig. S9).

We systematically tested the defined medium in more than 1300 experiments in order to properly assess all the components necessary for survival. We replaced these components with simpler building blocks in order to obtain a defined, minimal medium that contains only 26 components (19 of which are essential) (Fig. 1A). This medium, as predicted from our metabolic map and comparative analysis, also supports growth of *M. genitalium* (figs. S11 and S12). On the basis of these experiments, we estimated the upper flux limits for the use of the various nutrients (fig. S13). The medium implicitly validates the reconstructed metabolic map (Fig. 2), which consists of 189 reactions (table S2): 169 are catalyzed by the products of 140 known genes, and 20 are not yet assigned to any gene (table S4). The map includes 74 essential metabolic genes and 34 conditionally essential ones (depending on medium composition), which is in agreement with essentiality as determined by means of transposon mutagenesis analyses (with a 96% overlap) (fig. S14 and table S8) (*14*). A total of 32 enzymes (25%) are multifunctional; they have more than one activity and together catalyze 91 reactions (48% of the total) (table S3). With respect to previous genome annotations (*3*, *15*), we assigned new or refined functions to 57 metabolic genes (plus 30 nonmetabolic genes; see the new annotations in table S1). The above strategy could more generally be used to design media to grow axenically hard-to-culture bacteria, as was done for the recalcitrant *Tropheryma whipplei* (*16*) and might be applicable in the context of increasing metagenomics efforts.

Analysis of the metabolism of *M. pneumoniae* reveals that it is more linear than that of larger bacteria, such as *Bacillus subtilis* (Fig. 1B). Furthermore, *M. pneumoniae* has a wider metabolic network diameter (shortest biochemical pathway averaged over all pairs of substrates), although the diameter has been reported to increase with the logarithm of the network size (*17*). The

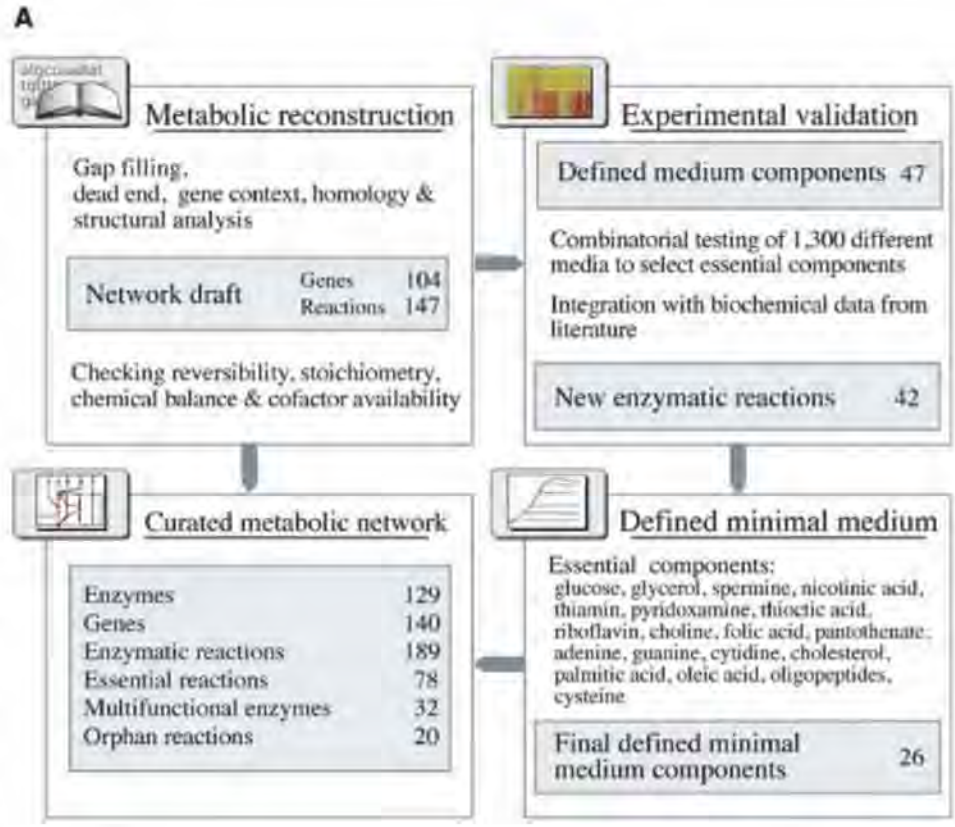
[1]Centre for Genomic Regulation (CRG) and Universitat Pompeu Fabra, Avenida Dr. Aiguader 88, 08003 Barcelona, Spain. [2]European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [3]Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127994, Russia. [4]Zentrum für Molekulare Biologie Heidelberg (ZMBH), Im Neuenheimer Feld 282, 69120 Heidelberg, Germany. [5]Institut Municipal d'Investigació Médica–Hospital del Mar, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Avenida Dr. Aiguader 88, 08003 Barcelona, Spain. [6]Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, Barcelona 08010, Spain.

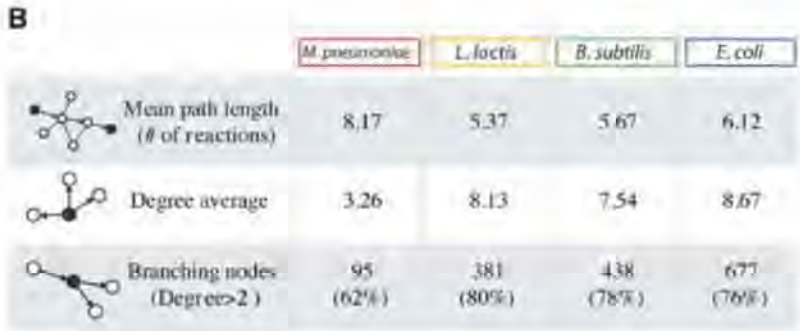
*To whom correspondence should be addressed. E-mail: luis.serrano@crg.es, bork@embl.de

greater linearity and the wider diameter of the network suggest that it is less interconnected and contains fewer parallel paths. Thus, the *M. pneumoniae* network is less redundant both in terms of enzyme paralogy and in network topology. Yet, the distribution of the number of metabolites per reaction is similar to other organisms (fig. S14). This is partly achieved by an increased fraction of multifunctional enzymes as compared with that in larger bacteria, as happens in endosymbionts (*17*). We did not find any evidence of *M. pneumoniae* multifunctional enzymes being more conserved than others. This suggests the larger number could be due to function acquisition that is not present (or detected) in their homologs. This might represent a more general mechanism expected to facilitate further genome reduction (Fig. 1, B and C). The increased linearity and limited redundancy in the metabolic network suggest limited robustness and adaptability to external factors (*18*): Of the metabolic enzymes, 60% are essential (*19*), in contrast to only 15% in *Escherichia coli* (www.shigen.nig.ac.jp/ecoli/pec/index.jsp).

M. pneumoniae has a relatively long duplication time (at least 8 hours) in comparison with *E. coli* or *L. lactis* (20 min), both in culture (*20*) and in the presence of host cells (*21*). Slow growth in genome-reduced, pathogenic bacteria has been proposed to be the result of (i) less efficient enzymatic activity that is explained by the accumulation of mutations resulting from genetic drift (*22*), (ii) a reduced number of ribosomal RNA (rRNA) operons, and/or (iii) other mechanisms related to the adaptation to a pathogenic lifestyle. To understand the causes of slow growth, it is necessary to measure the overall energetics of the metabolic network (Fig. 2) as well as the changes in macromolecules (Fig. 3A) and metabolites along the growth curve (Fig. 3B).

We used the metabolic map, the measured protein concentration (10 fg of protein per cell), and the estimated turnover rates of macromolecules (~20 hours for proteins and ~7 min for mRNA) (table S9 and fig. S15) to estimate the rate of glucose uptake required to duplicate a cell every 8 hours at 18,000 to 24,000 glucose molecules per second [assuming that the majority of adenosine triphosphate (ATP) is used for biomass production] (*10*). This figure closely matched the experimentally determined value under exponential growth: ~19,000 glucose molecules per cell per second (Fig. 3C) (*10*). When cultures approached stationary phase (Fig. 3A), the rate increased to ~45,000 glucose molecules per cell per second (Fig. 3C), concomitantly with the increased transcription of many glycolytic and fermentation genes (Fig. 3D and tables S10 and S11). In both cases, at least 95% of the glucose carbon was found in lactate and acetate (Fig. 3B and fig. S16), implying that the glucose is used primarily for energy production. At the fastest glucose consumption rate, assuming all ATP were devoted to biomass production, *M. pneumoniae* could divide about every





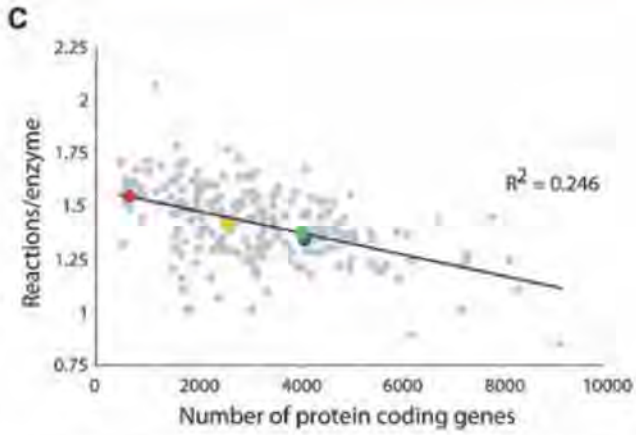


Fig. 1. Metabolic network development and properties and minimal medium design. (**A**) Schematic diagram of the process leading to *M. pneumoniae* metabolic network reconstruction and the design of a minimal medium. (**B**) Comparison of *M. pneumoniae* metabolic network properties with those of other model bacteria. (**C**) Quantification of enzyme multifunctionality among prokaryotic genomes. *M. pneumoniae*, red; *L. lactis*, yellow; *B. subtilis*, green; *E. coli*, blue; and other bacterial species, gray.

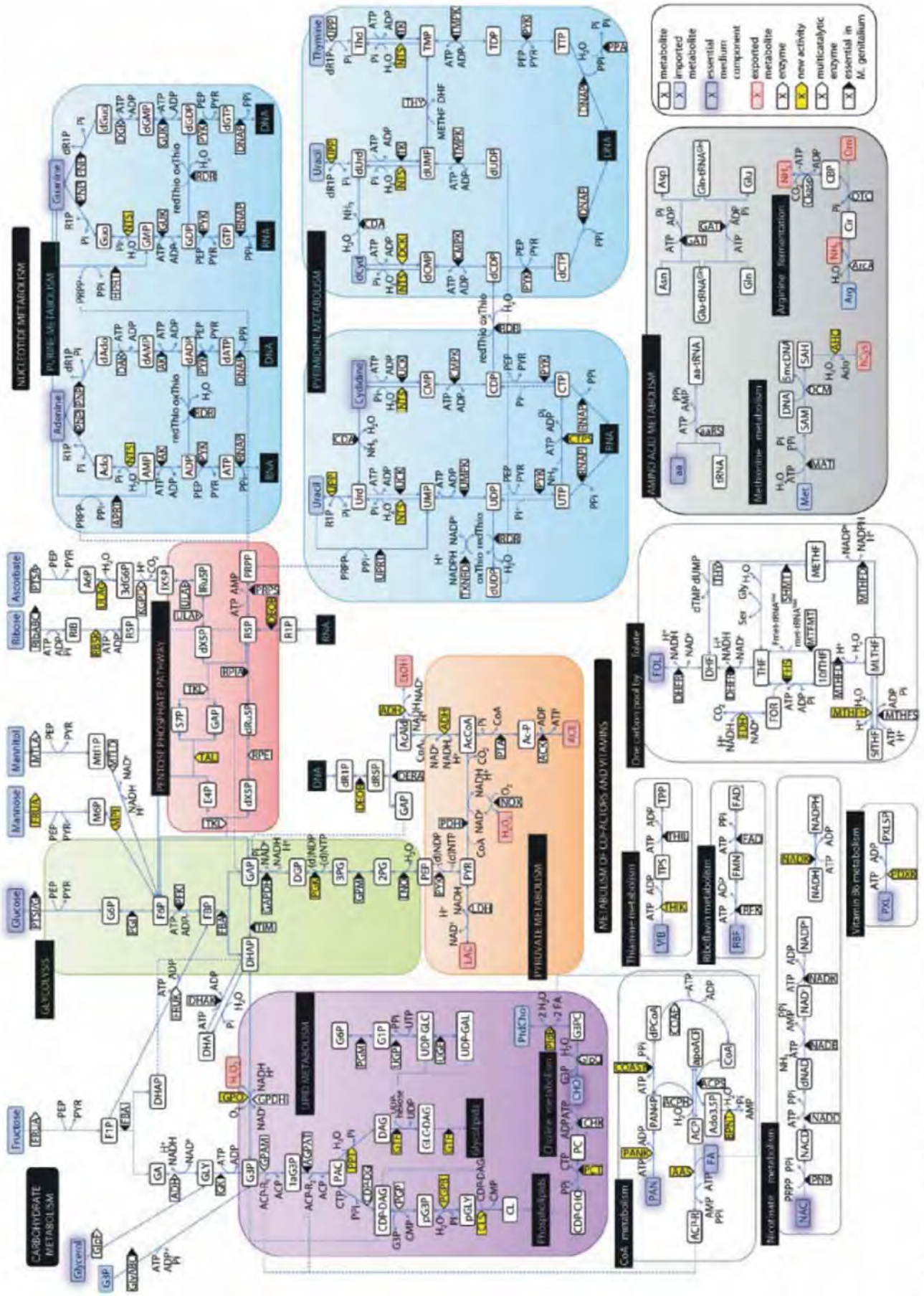


Fig. 2. Metabolic map of *M. pneumoniae*. Main metabolites are shown as boxes, and enzymes and transporters are shown as pentagons. Input metabolites are indicated in blue, and output products are indicated in red. New enzymatic activities determined in this study are displayed in yellow, and enzymes catalyzing multiple reactions are bold. Essential enzymes (according to the mutagenesis study in *M. genitalium*) are indicated with a black triangle. Minimal medium components have been shadowed in blue. See the bottom-right legend for details, fig. S12 and table S2 for description of the enzymatic reactions and enzymes, and table S25 for metabolite abbreviations. aaRS, aminoacyl-tRNA synthase.

3 hours. However, most of the energetic parameters (the concentration of glycolytic intermediates fructose-1,6-biphosphate, glycerol-3-phosphate, phosphoenolpyruvate, glucose-6-phosphate, fructose-6-phosphate, ribose-5-phosphate, and glycerone phosphate, as well as glucose uptake) that we measured were similar to those of larger bacteria (table S9) (*10*), which suggests comparable enzyme efficiencies. This similarity extended to regulatory processes seen in *Lactococcus lactis* (*23*). For example, as in *L. lactis*, we observed both a shift from mostly mixed-acid to homolactic fermentation and an acceleration of glycolysis when the medium acidifies (Fig. 3, A and B); the drop in O_2 concentration relieves inhibition of lactate dehydrogenase (*10*, *22*, *24*). Also, the ATP yield per fermented glucose (two to four ATP, depending on lactate or acetate fermentation) is the same as in *L. lactis* (table S9).

Given all of the above, we cannot explain the slow growth of *M. pneumoniae* on the basis of glycolytic efficiency or ATP yield. One of the main differences compared with fast dividing bacteria is the number of rRNA operons per genome [just one in *M. pneumoniae* and six in *L. lactis* (fig. S17) and five to 10 times proportionally fewer ribosomes as compared with those of *E. coli* (table S10)] (*7*). In many bacteria, the number of ribosomes correlates with the division rate (*25*). For *M. pneumoniae*, we see a correlation of changes in biomass duplication speed with the number of ribosomes but not with the glycolytic rate (Fig. 3, C and D, and fig. S17). We thus suggest that the slow division rate of *M. pneumoniae* is not due to less efficient energy production but to the limit in protein biosynthesis capacity. This small pathogenic bacterium does not appear to be optimized for biomass production. Instead, more complex strategies for fitness, such as suppression of growth by other microorganisms (*26*) or optimization of interactions with host cells, might determine growth rate in small organisms.

It has been suggested that genome-reduced organisms have limited adaptability to external factors (*24*). To determine the capacity of *M. pneumoniae* to respond to environmental changes, we performed three types of experiments. First, we followed the changes in gene expression from the exponential growth phase to the stationary phase (Fig. 4A). Analysis of changes in gene expression (validated by means of tiling arrays and quantitative polymerase chain reaction) (fig. S18) at different points along the growth curve showed that a large part of the transcriptome can be grouped into four time-dependent expression clusters (Fig. 4A, figs. S19 and S20, and tables S11 and S12). These clusters can be regarded as two pairs of anticorrelating patterns, indicating a complex regulation. Subsequent analysis by means of mass spectrometry for a subset of enzymes showed correlation between changes in mRNA and protein abundance (fig. S18 and table S10). For example, the production of lactate by lactate dehydrogenase (Mpn674|Ldh) revealed the close temporal coordination of gene and protein expression and metabolite turnover (Fig. 3, B and D, and table S10).

Second, we analyzed the response of *M. pneumoniae* to specific individual metabolic perturbations encountered as the population grows, such as low pH, accumulation of fermentation end-products, and sugar and amino acid starvation, as well as to more complex stimuli, such as entry into the stationary phase (Fig. 4, B to D, and tables S13 to S16). We found coordinated changes in gene expression specific to each condition (Fig. 4B and fig. S21). For example, there was a general inhibition of transcription and translation upon glucose deprivation and an increase of ATP proton pump genes at pH 6.5 (Fig. 4, B and C). Induction of the stringent response (a global response to the absence of amino acids) results in up-regulation of peptide and amino acid transporters (Fig. 4D). Also, a specific repression of the Thr-tRNA synthetase gene (*mpn553*) (table S17), which is a core component of a tRNA synthetase complex (*7*), suggests its possible regulatory role in complex assembly and therefore in regulation of translation. We found some common responses to multiple stresses. Some were known, such as the

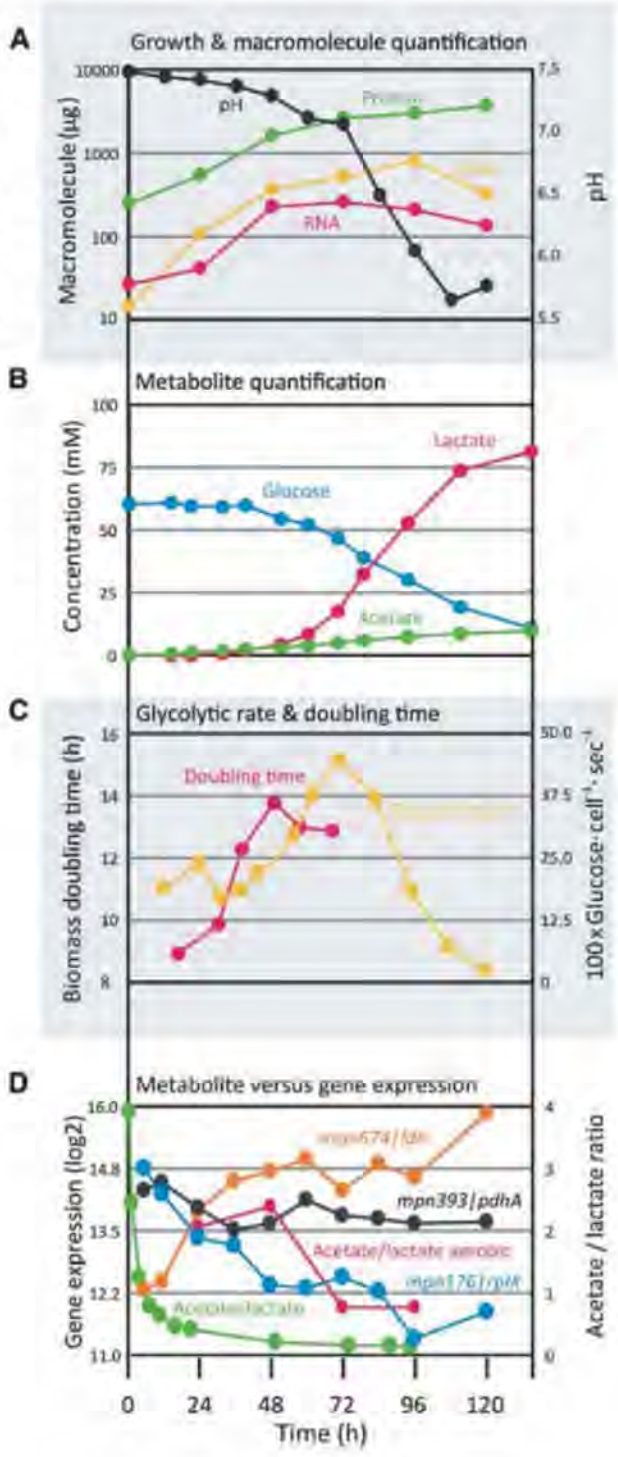


Fig. 3. Determination of various metabolic parameters in growing cultures. Consistently generated heterogeneous data, all derived by using a rich medium, are compared through time in hours (*x* axis). (**A**) *M. pneumoniae* growth determined by monitoring the decrease in extracellular pH and the concomitant changes in the amount of protein, DNA, and total RNA. (**B**) Determination of glucose consumption and its fermentation to lactate and acetate. (**C**) Changes in the number of glucose molecules imported by a cell per second and comparison with the biomass doubling time. (**D**) Changes in gene expression of a representative ribosomal protein (*rplX*) and two enzymes [*ldh* and *pdhA* (a component of the pyruvate dehydrogenase complex), which are enzymes from the two fermentation branches] and the relation with the shift from acetate to lactate production (the ratio between acetate and lactate is shown in green, which can be compared with that of cells grown in the presence of oxygen, shown in red).

down-regulation of ribosomal proteins or peptide importers, which is common to all stresses. Others, like the up-regulation of *ldh* and glycerol-3-P dehydrogenase (*mpn051*), were unexpected and suggest additional functions for these proteins during stress (Fig. 4B).

Third, we adapted the cells by means of serial passage (15 passages) to efficient growth in other carbon sources (fructose, mannose, and glycerol) (tables S18 to S20). Fructose adaptation resulted in overexpression of fruA and fruK (>3 $\log_2$), and mannose-adapted cells overexpressed the mannitol importer (>5 $\log_2$) (tables S19 and S20). Thus, *M. pneumoniae* shows surprising adaptation capability similar to that reported for *E. coli* (*27*).

The coordinated changes in gene expression along the growth curve, the specific responses to many various metabolic perturbations, and the adaptability of the cells to various carbon sources indicate that *M. pneumoniae* retains some robustness and adaptability despite its extreme genome reduction.

Compared with more complex bacteria, *M. pneumoniae* lacks the majority of transcription factors (TFs) regulating metabolic gene factors [such as the catabolite regulation protein (CRP)], major sigma factors, and other regulators (*28*). Gene assignment on the basis of sequence analysis (table S1), in some cases validated through copurification with the RNA polymerase complex (such as *mpn266|spxA*) (*7*), revealed four TFs (*mpn239|gntR*, *mpn329|fur*, and *mpn124|hrcA*), the general sigma 70 factor (*mpn352|sigA*), two putative sigma-like factors (*mpn626|sigD* and *mpn424|ylxM*), and a putative DNA-binding protein (*mpn241|whiA*) (fig. S2 and table S1). Despite this apparently reduced gene regulatory toolbox, both environmental stresses (*6*) and metabolic insults induced complex, specific transcriptional responses; comparison with more complex bacteria showed sim-

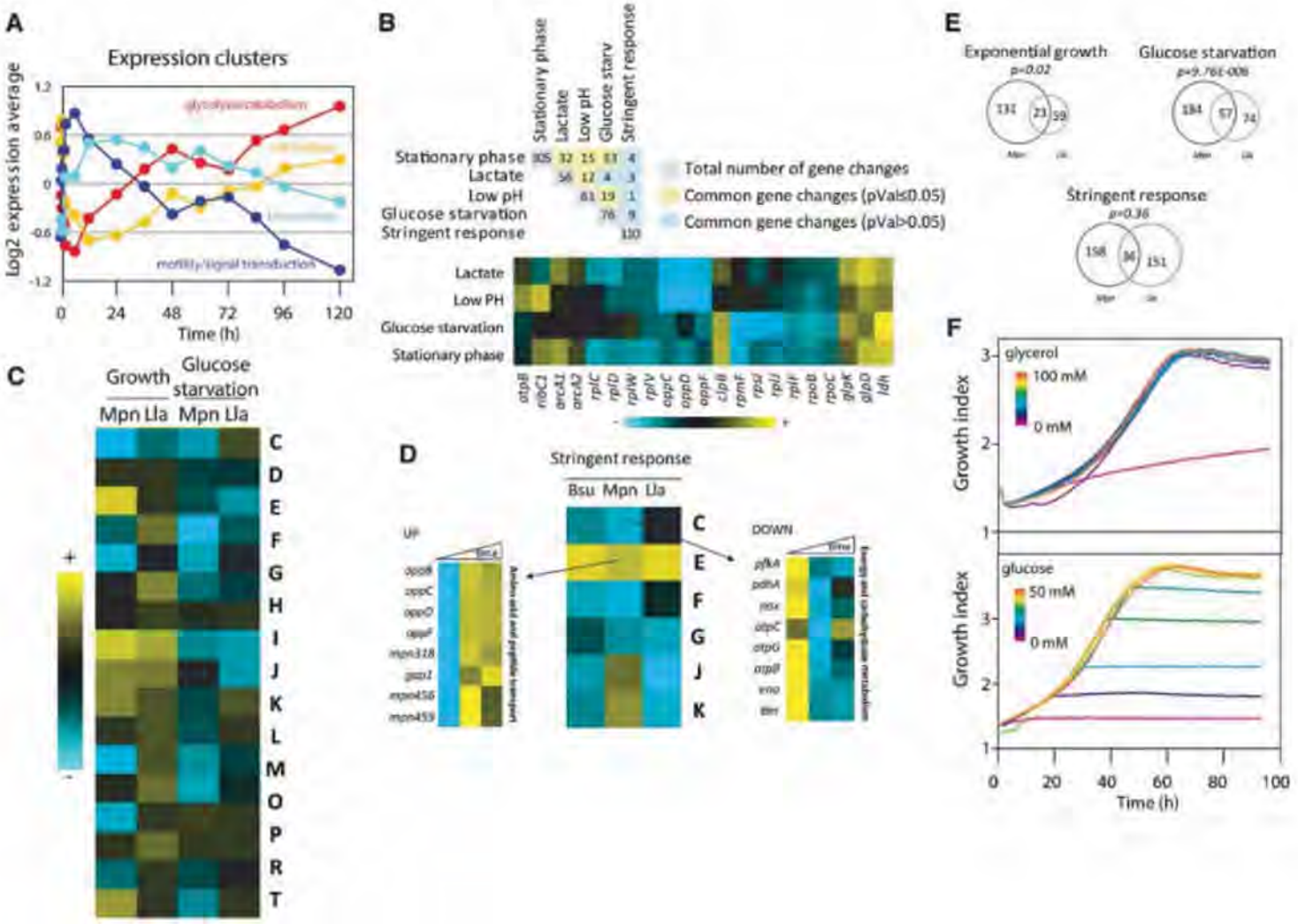


Fig. 4. Regulation of metabolism. (**A**) Representative plot of the four main gene co-expression clusters identified along the growth curve, named after the main functional classes of the genes involved. (**B**)(Top) Overlap between changes in gene expression under various stresses: lactate (80 mM buffered lactate), low pH (pH 6.5), glucose or amino acid starvation (stringent response), and the entry into stationary phase. (Bottom) Heat map of the genes found to be commonly up- or down-regulated under stress and growth inhibition. (**C**) Gene Ontology functional classification of genes significantly regulated during exponential growth and glucose deprivation in *M. pneumoniae* (*Mpn*) or *L. lactis* (*Lla*). The average of the significant changes within each category is depicted. C, energy production and conversion; D, cell division and chromosome partitioning; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme metabolism; I, lipid metabolism; J, translation, ribosomal structure, and biogenesis; K, transcription; L, DNA replication, recombination, and repair; M, cell envelope biogenesis, outer membrane; O, posttranslational modification, protein turnover, and chaperones; P, inorganic ion transport and metabolism; R, general function; T, signal transduction mechanisms. (**D**) Stringent response expression pattern was compared with that of *L. lactis* and *B. subtilis* (Bsu). The average of the significant changes upon stringent response induction (with norvaline) is shown. (**E**) Venn diagrams showing the overlap in *M. pneumoniae* and *L. lactis* of ortholog genes under various metabolic conditions. High *P* value in the case of the stringent response indicates that it is not statistically significant. (**F**) (Top) Growth curve of cells growing in minimal medium plus increasing amounts of glycerol. (Bottom) Glucose titration in Hayflick is shown for comparison.

ilarities but also some specific differences in regulation of gene expression (tables S21 to S23). For example, we observed an increase in mRNA and protein expression of glycolytic enzymes concomitant with the increase of glycolytic rate upon medium acidification (Fig. 3, C and D, fig. S20, and table S15), very similar to what has been described in *L. lactis* cultures (*29*). Response to glucose starvation was also similar to that of *L. lactis* (Fig. 4, C and E, and table S21) (*30*). Part of the stringent response, such as the induction of peptide and amino acid transporters and down-regulation of carbohydrate catabolism (*31*), was conserved in *M. pneumoniae* (table S22); other mechanisms, such as the repression of ribosomal protein operons or rRNA synthesis, were not observed (Fig. 4D and fig. S22). This is in agreement with the proposed involvement of the RNA polymerase omega subunit (missing in *M. pneumoniae*) in sensing guanosine pentaphosphate/tetraphosphate [(p)ppGpp] and thus arresting rRNA biosynthesis (*32*).

We believe it unlikely that the conserved responses, and the specific differences in regulation, can be caused only by combinations of the few TFs that regulate operons and suboperons, even if one includes regulation through antisense RNA (*6*). The presence of genes for synthesis or degradation of a number of chemical messengers, such as (p)ppGpp (*mpn397|spoT*), AppppA (*mpn273|hit1*), or c-di-AMP (*mpn244|disA*) (fig. S2 and table S1) (*33*), implies that signaling mechanisms have been preserved despite genome reduction. For example, overexpression of the spoT gene that regulates (p)ppGpp levels (*31*) results in substantial changes in gene expression, mainly related to the stringent response (table S24). The presence of genes coding for a Ser/Thr phosphatase (*mpn247|ptc1*), two protein kinases (Ser/Thr/Tyr kinase *mpn248|prkC* and *mpn223|hrpK*, an HPr kinase), and the differential phosphorylation of key metabolic enzymes under various growth conditions (*33*) suggest posttranslational control. Also, metabolites such as glycerol regulate gene expression at the base of the fermentation branches in *M. pneumoniae* (*34*) as well as glucose import (*35*). This explains why glycerol is essential in the minimal medium in a concentration-independent manner (Fig. 4F).

Our results suggest that complex metabolic regulation can be achieved in a streamlined genome despite the absence of the respective TFs probably because of a combination of transcriptional regulators, posttranslational modifications, and small molecules, including chemical messengers and metabolites.

Taken together, our newly established *M. pneumoniae* resource, containing a manually annotated metabolic map, full annotations, reactome, consistently measured growth curves, and gene expression profiles corresponding to an extensive list of metabolites, should facilitate integrative systems biology studies at a high resolution. Comparison with more complex bacteria revealed systemic features associated with genome streamlining, which should be examined in other small bacteria. Despite its apparent simplicity, we have shown that *M. pneumoniae* shows metabolic responses and adaptation similar to more complex bacteria, providing hints that other, unknown regulatory mechanisms might exist.

References and Notes

1. A. M. Feist, B. O. Palsson, *Nat. Biotechnol.* **26**, 659 (2008).
2. K. B. Waites, D. F. Talkington, *Clin. Microbiol. Rev.* **17**, 697 (2004).
3. T. Dandekar *et al.*, *Nucleic Acids Res.* **28**, 3278 (2000).
4. T. Proft, R. Herrmann, *Mol. Microbiol.* **13**, 337 (1994).
5. E. van Nimwegen, *Trends Genet.* **19**, 479 (2003).
6. M. Güell *et al.*, *Science* **326**, 1268 (2009).
7. S. Kühner *et al.*, *Science* **326**, 1235 (2009).
8. J. D. Pollack, *Trends Microbiol.* **5**, 413 (1997).
9. M. Pachkov, T. Dandekar, J. Korbel, P. Bork, S. Schuster, *Gene* **396**, 215 (2007).
10. Materials and methods are available as supporting material on *Science* Online.
11. J. D. Pollack, M. V. Williams, R. N. McElhaney, *Crit. Rev. Microbiol.* **23**, 269 (1997).
12. S. Halbedel, C. Hames, J. Stulke, *J. Bacteriol.* **186**, 7936 (2004).
13. J. T. Manolukas, M. F. Barile, D. K. Chandler, J. D. Pollack, *J. Gen. Microbiol.* **134**, 791 (1988).
14. C. A. Hutchison *et al.*, *Science* **286**, 2165 (1999).
15. R. Himmelreich *et al.*, *Nucleic Acids Res.* **24**, 4420 (1996).
16. P. Renesto *et al.*, *Lancet* **362**, 447 (2003).
17. E. Zientz, T. Dandekar, R. Gross, *Microbiol. Mol. Biol. Rev.* **68**, 745 (2004).
18. L. Z. Pipe, M. J. Grimson, *Mol. Biosyst.* **4**, 192 (2008).
19. J. I. Glass *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 425 (2006).
20. B. M. Hasselbring, J. L. Jordan, R. W. Krause, D. C. Krause, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16478 (2006).
21. S. F. Dallo, J. B. Baseman, *Microb. Pathog.* **29**, 301 (2000).
22. J. B. Smart, T. D. Thomas, *Appl. Environ. Microbiol.* **53**, 533 (1987).
23. N. A. Moran, *Cell* **108**, 583 (2002).
24. M. Cocaign-Bousquet, S. Even, N. D. Lindley, P. Loubiere, *Appl. Microbiol. Biotechnol.* **60**, 24 (2002).
25. M. Nomura, *J. Bacteriol.* **181**, 6857 (1999).
26. B. Teusink *et al.*, *J. Biol. Chem.* **281**, 40041 (2006).
27. E. Oxman, U. Alon, E. Dekel, *Evolution* **62**, 1547 (2008).
28. W. Goebel, S. Lory, *Curr. Opin. Microbiol.* **9**, 123 (2006).
29. S. Even, N. D. Lindley, M. Cocaign-Bousquet, *Microbiology* **149**, 1935 (2003).
30. S. Even, N. D. Lindley, P. Loubiere, M. Cocaign-Bousquet, *Mol. Microbiol.* **45**, 1143 (2002).
31. K. Potrykus, M. Cashel, *Annu. Rev. Microbiol.* **62**, 35 (2008).
32. C. E. Vrentas, T. Gaal, W. Ross, R. H. Ebright, R. L. Gourse, *Genes Dev.* **19**, 2378 (2005).
33. H. C. Su, C. A. Hutchison 3rd, M. C. Giddings, *BMC Microbiol.* **7**, 63 (2007).
34. S. Halbedel *et al.*, *J. Mol. Biol.* **371**, 596 (2007).
35. S. Halbedel, C. Hames, J. Stulke, *J. Mol. Microbiol. Biotechnol.* **12**, 147 (2007).

36. The authors thank V. Benes and S. Schmidt of the Gene Core facility at EMBL (Heidelberg, Germany) for hybridization of samples in custom made arrays; A. Wieslander (Stockholm University, Sweden) for advice on the lipid requirements for the minimal medium; J. Marcos del Aguila (Pompeu Fabra University, Barcelona, Spain) for conceptual and experimental input on the gas chromatography–mass spectrometry measurements; T. Doerks (EMBL, Heidelberg, Germany) for operon analysis and help in annotation; and J. Leigh-Bell for editorial help. L.S. is an Institució Catalana de Recerca i Estudis Avançats professor.

Supporting online Material

www.sciencemag.org/cgi/content/full/326/5957/1263/DC1
Materials and Methods
SOM Text
Figures S1 to S22
Tables S1 to S25
References

3 June 2009; accepted 2 October 2009
10.1126/science.1177263

Transcriptome Complexity in a Genome-Reduced Bacterium

Marc Güell,[1] Vera van Noort,[2] Eva Yus,[1] Wei-Hua Chen,[2] Justine Leigh-Bell,[1] Konstantinos Michalodimitrakis,[1] Takuji Yamada,[2] Manimozhiyan Arumugam,[2] Tobias Doerks,[2] Sebastian Kühner,[2] Michaela Rode,[2] Mikita Suyama,[2]* Sabine Schmidt,[2] Anne-Claude Gavin,[2] Peer Bork,[2]† Luis Serrano[1,3]†

To study basic principles of transcriptome organization in bacteria, we analyzed one of the smallest self-replicating organisms, *Mycoplasma pneumoniae*. We combined strand-specific tiling arrays, complemented by transcriptome sequencing, with more than 252 spotted arrays. We detected 117 previously undescribed, mostly noncoding transcripts, 89 of them in antisense configuration to known genes. We identified 341 operons, of which 139 are polycistronic; almost half of the latter show decaying expression in a staircase-like manner. Under various conditions, operons could be divided into 447 smaller transcriptional units, resulting in many alternative transcripts. Frequent antisense transcripts, alternative transcripts, and multiple regulators per gene imply a highly dynamic transcriptome, more similar to that of eukaryotes than previously thought.

Although large-scale gene expression studies have been reported for various bacteria (*1*–*7*), comprehensive strand-specific data sets are still missing, limiting our understanding of operon structure and regulation. Similarly, the number of classified noncoding RNAs in bacteria has recently been expanded (*8*), but a complete and unbiased repertoire is still not available. To obtain a blueprint of bacterial transcription, we combined the robustness and versatility of spotted arrays [62 independent conditions and 252 array experiments (*9*)], the superior res-

olution of strand-specific tiling arrays (Fig. 1A) (designed after genome resequencing, table S1), and the mapping capacity of RNA deep sequencing [direct strand-specific sequencing (DSSS)] (Fig. 1A and fig. S1) to analyze one of the smallest bacteria that can live outside a host cell, *Mycoplasma pneumoniae*, with 689 annotated protein-coding genes and 44 noncoding RNAs (ncRNAs).

Considering DSSS under reference conditions (*9*) and 43 tiling arrays from four time series (growth curve, heat shock, DNA damage, and cell cycle arrest) (table S8), we observed the expression of all genes. Using a segmentation algorithm for the tiling arrays (*10*), we identified an additional 117 regions with no previous annotation (table S2) (*9*). These regions were further confirmed by DSSS (Fig. 1B and fig. S1) and in four cases by quantitative polymerase chain reaction (table S3). Sequence similarity with known proteins revealed the presence of two previously undescribed protein-coding genes, one pseudogene, one N-terminal truncation, and five 5′ extensions of known genes (table S2). The remaining 108 transcripts are probably regulatory rather than structural RNAs, because comparison of their predicted secondary structures with the ones of coding genes does not show any substantial difference (*9*). Eighty-nine of them are antisense with respect to previously annotated genes. Out of the nonoverlapping ones, two of them (NEW87 and NEW8) are conserved in *M. genitalium* and could be involved in DNA replication and repair, and in peptide transport, respectively (*9*) (figs. S3 to S5). In total, 13% of the coding genes are covered by antisense; this is twice more than in yeast (7%) (*11*) and about half of what was reported for plants (22.2%) (*12*, *13*) or humans (22.6%) (*14*). Antisense transcripts may affect expression of the overlapping functional sense transcripts through several mechanisms (*15*): Double-stranded RNA-dependent mechanisms require coexpression with their targets (*16*), whereas transcriptional interference implies the mutual exclusion of sense and antisense transcripts (*17*, *18*). In *M. pneumoniae*, we observed a predominance of double-stranded RNA mechanisms as in mammals (*19*) (47% positive correlation versus 2% negative correlation). In addition, we detected a reduced expression level of genes targeted by antisense transcripts, as reported in some prokaryotes (*9*, *18*) (fig. S6).

We identified operon boundaries through sharp transcription changes in the tiling reference condition by using local convolution methods (Fig. 1A) (*9*, *20*). More than 90% of the operons (139 polycistronic and 202 monocistronic operons, table S4) were well supported by DSSS reads [DSSS alone was not sufficient to unambiguously characterize operons (fig. S2) (*9*)]. Most polycistronic operons contain two or three genes (Fig. 1C, fig. S7, and table S4); the largest one is the ribosomal operon containing 20 genes. For the majority of operons, we observed a canonical or slightly altered version of a standard sigma 70 promoter region (fig. S8), with transcription starts located within 60 base pairs (fig. S9) upstream of the translation start (*6*). In contrast to previous suggestions (*21*), we observed, as proposed by others (*22*), a preferential use of termination hairpins for the tight regulation of gene expression (Fig. 1, A and D, and table S5). Moreover, we found that almost half of the consecutive genes within polycistronic operons show a decay behavior (Fig. 1A and fig. S1), indicating that such staircase-like expression is a widespread phenomenon in bacteria (*9*).

Analysis of the 43 tiling arrays and integration with 252 spotted arrays representing 173 independent conditions, some of them from time series, revealed context-dependent modulation of operon structures involving repression or activation of operon internal genes, as well as of genes located at the beginning or end (Fig. 2, A and B, fig. S10, and table S5). In some cases this modulation can be assigned to specific environmental changes. Down-regulation of the first four genes of the *ftsZ* operon involved in the initiation of cell division corresponds to entry into stationary phase (Fig.

[1]Centre for Genomic Regulation (CRG), Universitat Pompeu Fabra, Barcelona, Spain. [2]European Molecular Biology Laboratory (EMBL), Heidelberg, Germany. [3]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain.

*Present address: Center for Genomic Medicine, Graduate School of Medicine, University of Kyoto, Kyoto, Japan.
†To whom correspondence should be addressed. E-mail: bork@embl.de (P.B.); luis.serrano@crg.es (L.S.)

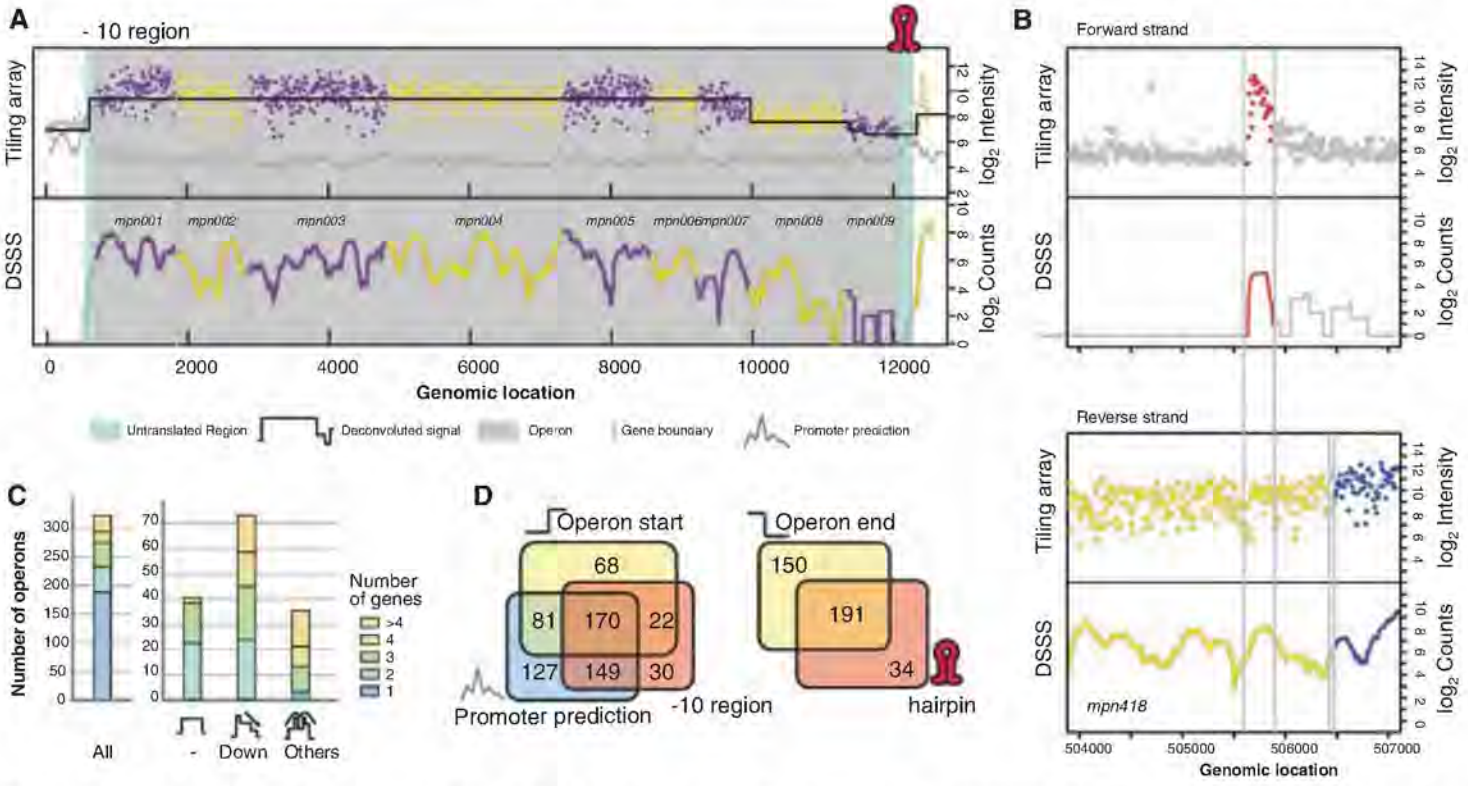
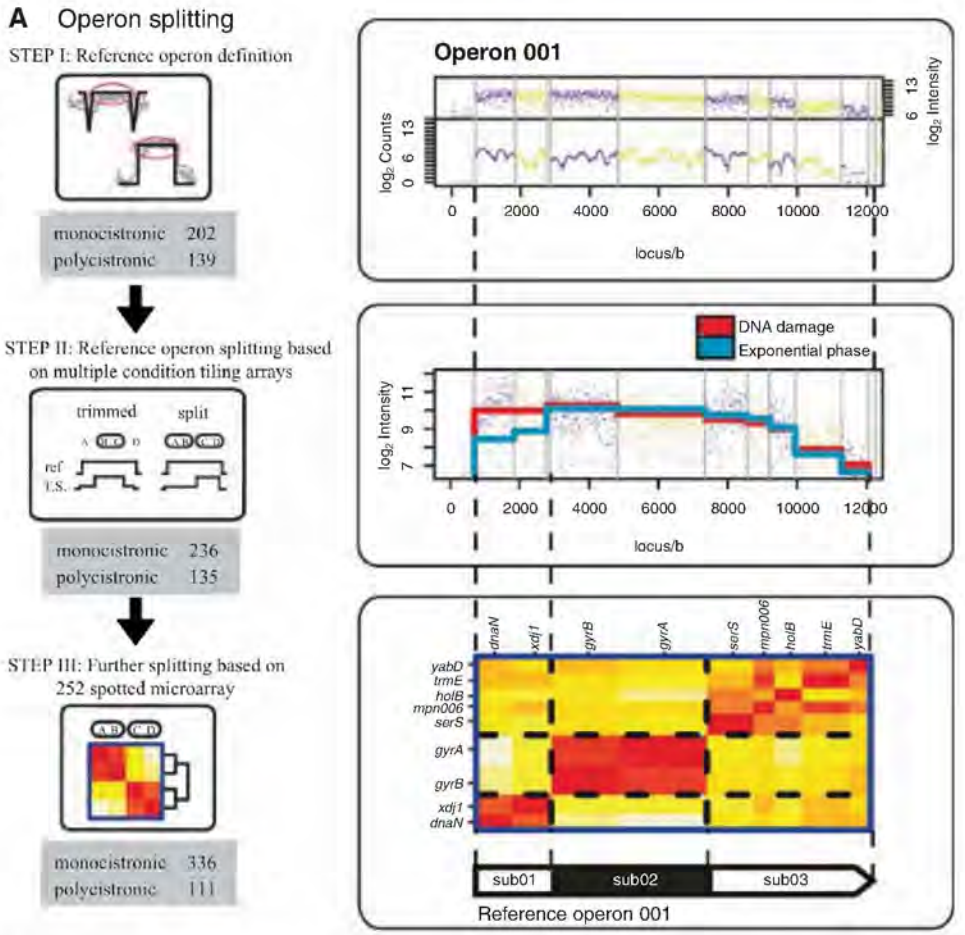
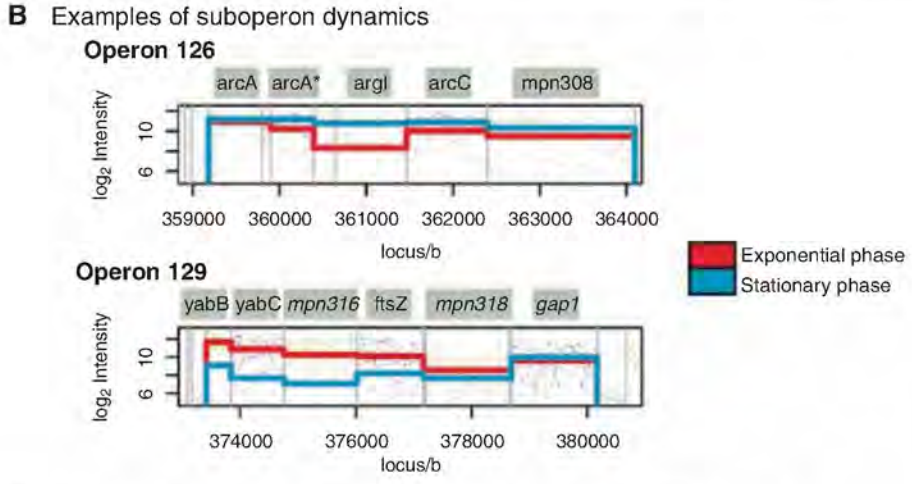
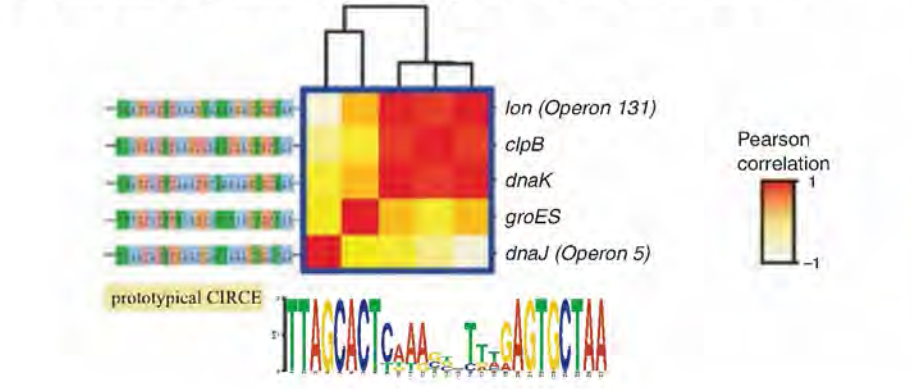


Fig. 1. Transcriptome feature in the reference condition. (**A**) The first operon in the genome on the forward strand has a staircase behavior, meaning that the consecutive genes have lower and steady expression levels. (**B**) Example of an antisense RNA transcript. (**C**) Analysis of staircase operons. (Left) All reference operons subdivided by the number of protein-coding genes they contain. (Right) All reference operons subdivided by their staircase behavior (see bottom graphs). (**D**) (Left) Overlap of operon starts and single-gene starts with previously identified −10 promoter sequence motifs in *M. pneumoniae* (*29*) and predicted promoters based on hexamers. (Right) Overlap of operon ends and single-gene ends with predicted transcription termination hairpins.

Fig. 2. Alternative operon structure. The continuous lines in (A) and (B) indicate expression level measured with tiling arrays. (**A**) Alternative transcripts discovery pipeline (*9*). Reference operon 001 is split into three suboperons. (Top) Tiling and DSSS under reference conditions. (Middle) Specific expression changes for genes **dNAa** and *xdj1* involved in DNA repair and replication. (Bottom) The coexpression matrices correspond to the final conditional operon splitting by 252 arrays. (**B**) Two examples of conditional operons are presented. (Top) Specific induction of the middle genes in operon 126 when the cells reach stationary phase. (Bottom) Repression of the first four genes of the operon 129 involved in cell division, when the cells reach stationary phase. (**C**), Example of heat shock–induced genes sharing the known CIRCE element. The calculated consensus sequence is represented below.







2B, lower panel). An increase in the expression of arginine fermentation genes (*arcA*, *arcI*, and *arcC*) (Fig. 2B) in stationary phase could be a mechanism to cope with acidification (*23*). We found formal evidence for a total of 447 transcriptional units (336 monocistronic and 111 polycistronic), implying a high rate of alternative transcripts (42%) in this bacterium under the conditions studied, similar to that in eukaryotes (40%, although still under debate) (*24*) and archea (40% in *H. salinarum*) (*25*). We found that genes that are split into different suboperons tend to belong to different functional categories (*9*). Thus, although genome reduction leads to longer operons accommodating genes with different functions (*26*), the latter can still retain internal transcription and termination sites under certain conditions.

The high frequency of alternative transcripts of *M. pneumoniae* genes hints at a situation similar to that in eukaryotes, where many factors contribute to the regulation of gene expression. To further support this hypothesis, we used gene expression clustering under the 62 distinct conditions (table S7) to identify groups of coexpressed genes and their possible common regulatory motifs. Using a correlation cutoff of 0.65, we identified 94 coexpression groups (table S6 and fig. S11), encompassing 416 genes. Thirty of the clusters contained genes from more than two operons. Of these, 14 share a specific sequence motif in their upstream region and another 8 have a specific combination of motifs (fig. S12), which might drive the coexpression (for example, 4 of the 14 motifs are found at splitting sites inside operons). The rest of the genes did not group together, implying complex and multiple levels of regulation orchestrated by the various environmental conditions. This is exemplified by the five heat shock–induced genes containing a regulatory CIRCE (controlling inverted repeat of chaperone expression) element (*27*) (Fig. 2C). Not all of them clustered together, indicating at least one other regulatory element. Similarly, overexpression of a transcription factor (Fur, ferric uptake regulator) reveals a common motif in all genes significantly changing expression, although they belong to different coexpression clusters (fig. S13 and table S6).

Our work revealed an unanticipated complexity in the transcriptome of a genome-reduced bacterium. This complexity cannot be explained by the presence of eight predicted transcription factors (*26*). Furthermore, the fact that the proteome organization is not explainable by the genome organization (*28*) indicates the existence of other regulatory processes. The surprisingly frequent expression heterogeneity within operons, the change of operon structures leading to alternative transcripts in response to environmental perturbations, and the frequency of antisense RNA, which might explain some of these expression changes, suggest that transcriptional regulation in bacteria resemble that of eukaryotes more than previously thought.

References and Notes

1. B. Tjaden *et al.*, *Nucleic Acids Res.* **30**, 3732 (2002).
2. D. W. Selinger *et al.*, *Nat. Biotechnol.* **18**, 1262 (2000).
3. N. B. Reppas, J. T. Wade, G. M. Church, K. Struhl, *Mol. Cell* **24**, 747 (2006).
4. C. M. Nelson *et al.*, *BMC Genomics* **9**, 364 (2008).
5. T. Akama *et al.*, *J. Bacteriol.* **191**, 3321 (2009).
6. P. T. McGrath *et al.*, *Nat. Biotechnol.* **25**, 584 (2007).
7. A. Toledo-Arana *et al.*, *Nature* **459**, 950 (2009).
8. J. Vogel, E. G. Wagner, *Curr. Opin. Microbiol.* **10**, 262 (2007).
9. Materials and methods are available as supporting material on *Science* Online.
10. W. Huber, J. Toedling, L. M. Steinmetz, *Bioinformatics* **22**, 1963 (2006).
11. Z. Xu *et al.*, *Nature* **457**, 1033 (2009).
12. X. J. Wang, T. Gaasterland, N. H. Chua, *Genome Biol.* **6**, R30 (2005).
13. S. R. Henz *et al.*, *Plant Physiol.* **144**, 1247 (2007).
14. X. Ge, Q. Wu, Y. C. Jung, J. Chen, S. M. Wang, *Bioinformatics* **22**, 2475 (2006).
15. M. Lapidot, Y. Pilpel, *EMBO Rep.* **7**, 1216 (2006).
16. S. Brantl, E. G. Wagner, *EMBO J.* **13**, 3599 (1994).
17. G. Andre *et al.*, *Nucleic Acids Res.* **36**, 5955 (2008).
18. S. Brantl, *Curr. Opin. Microbiol.* **10**, 102 (2007).
19. S. Katayama *et al.*, *Science* **309**, 1564 (2005).
20. S. D. Hooper *et al.*, *Mol. Syst. Biol.* **3**, 72 (2007).
21. T. Washio, J. Sasayama, M. Tomita, *Nucleic Acids Res.* **26**, 5456 (1998).
22. M. J. de Hoon, Y. Makita, K. Nakai, S. Miyano, *PLOS Comput. Biol.* **1**, e25 (2005).
23. A. Budin-Verneuila, E. Maguin, Y. Auffraya, S. Dusko Ehrlich, V. Pichereaua, *Lait* **84**, 8 (2004).
24. S. Boue, I. Letunic, P. Bork, *Bioessays* **25**, 1031 (2003).
25. T. Koide *et al.*, *Mol. Syst. Biol.* **5**, 285 (2009).
26. E. Yus *et al.*, *Science* **326**, 1263 (2009).
27. L. J. Chang, W. H. Chen, F. C. Minion, D. Shiuan, *Biochem. Biophys. Res. Commun.* **367**, 213 (2008).
28. S. Kühner *et al.*, *Science* **326**, 1235 (2009).
29. J. Weiner III, R. Herrmann, G. F. Browning, *Nucleic Acids Res.* **28**, 4488 (2000).

30. We thank the Genomics core facility at EMBL (Heidelberg, Germany), J. Lozano for help with statistical analysis, and the Ultrasequencing Unit at CRG. This work was funded by the Foundation Marcelino Botín, the Ministry of Education of Spain (MEC)–Consolider, and the European Research Council. M.G. is funded by the Spanish MEC–Formación Profesorado Universitario. V.N. is funded by the Netherlands Organization for Scientific Research.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1268/DC1
Materials and Methods
SOM Text
Figs. S1 to S14
Tables S1 to S8
References

28 May 2009; accepted 2 October 2009
10.1126/science.1176951

Crystal Structure of the Catalytic Core of an RNA-Polymerase Ribozyme

David M. Shechner,[1,2] Robert A. Grant,[2] Sarah C. Bagby,[1,2] Yelena Koldobskaya,[3] Joseph A. Piccirilli,[3] David P. Bartel[1,2]*

Primordial organisms of the putative RNA world would have required polymerase ribozymes able to replicate RNA. Known ribozymes with polymerase activity best approximating that needed for RNA replication contain at their catalytic core the class I RNA ligase, an artificial ribozyme with a catalytic rate among the fastest of known ribozymes. Here we present the 3.0 angstrom crystal structure of this ligase. The architecture resembles a tripod, its three legs converging near the ligation junction. Interacting with this tripod scaffold through a series of 10 minor-groove interactions (including two A-minor triads) is the unpaired segment that contributes to and organizes the active site. A cytosine nucleobase and two backbone phosphates abut the ligation junction; their location suggests a model for catalysis resembling that of proteinaceous polymerases.

The RNA world hypothesis proposes that early life forms lacked DNA and coded proteins, depending instead on RNA for both chemical catalysis and information storage (*1*). Central to this RNA world would have been polymerase ribozymes able to replicate RNA. Among the efforts to generate ribozymes with such ability, the most productive have started with the class I RNA ligase ribozyme (*2*–*4*). This ribozyme was originally isolated from a large pool of random sequences (*5*, *6*). It has since been improved by mutation and selection and has served as a platform for modeling ribozyme evolution in vitro (*6*–*8*). Because it rapidly promotes a reaction with chemistry identical to that catalyzed by proteinaceous enzymes that replicate RNA (Fig. 1A) (*6*), the ligase has provided the catalytic engine for more sophisticated RNA enzymes that use nucleoside triphosphates and the information from an external RNA template to synthesize short strands of RNA (*2*, *3*, *9*). Although more efficient with some templates than with others, this primer-extension reaction is gen-

[1]Whitehead Institute for Biomedical Research and Howard Hughes Medical Institute, 9 Cambridge Center, Cambridge, MA 02142, USA. [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Department of Chemistry, Department of Biochemistry and Molecular Biology, and Howard Hughes Medical Institute, University of Chicago, Chicago, IL 60637, USA.

*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

eral in that all templates tested support detectable extension (*2*–*4*). To understand the structural basis behind RNA-catalyzed RNA polymerization, we have solved the crystal structure of the class I ligase ribozyme.

The ligase sequence variant we crystallized was the product of three successive in vitro selection experiments, the last of which mutagenized segments not participating in known base pairs (termed "joining regions") and selected variants that folded and reacted within milliseconds (*5*, *6*, *10*). This experiment produced an improved variant that, unlike its predecessor, yielded useful crystals (data to 3.0 Å, tables S1 to S3, figs. S1 and S2) (*11*). This variant is more tolerant of low Mg^{2+} concentrations; it reacts 15 times faster than the predecessor in 1 mM Mg^{2+} (*10*) but only slightly faster than the predecessor in high Mg^{2+} [predecessor reaction rate, 800/min in 60 mM Mg^{2+}, pH 9 (*12*)]. As with the predecessor, its reaction is pH-dependent, slowing to

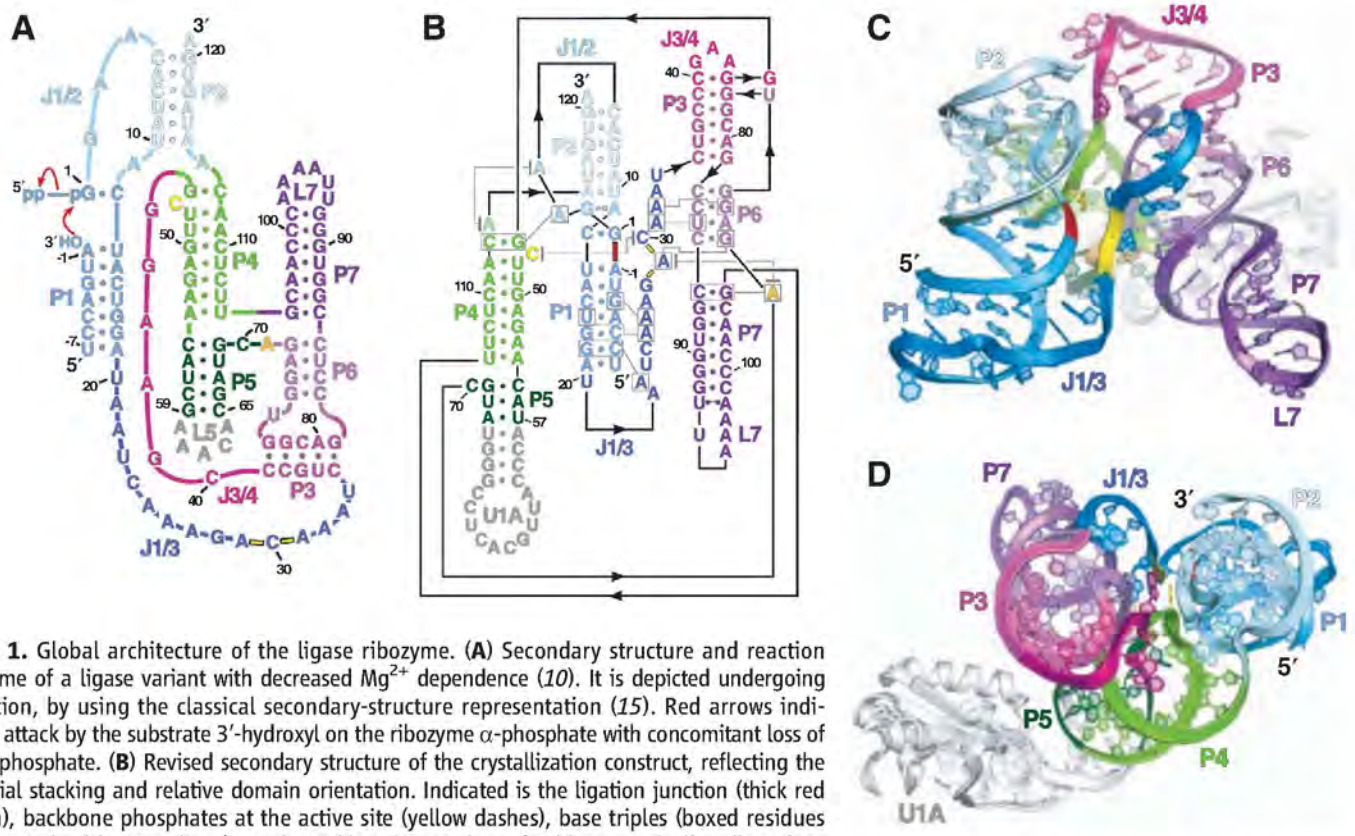


Fig. 1. Global architecture of the ligase ribozyme. (**A**) Secondary structure and reaction scheme of a ligase variant with decreased Mg^{2+} dependence (*10*). It is depicted undergoing ligation, by using the classical secondary-structure representation (*15*). Red arrows indicate attack by the substrate 3′-hydroxyl on the ribozyme α-phosphate with concomitant loss of pyrophosphate. (**B**) Revised secondary structure of the crystallization construct, reflecting the coaxial stacking and relative domain orientation. Indicated is the ligation junction (thick red dash), backbone phosphates at the active site (yellow dashes), base triples (boxed residues connected with gray lines), and stacking interactions (residues vertically aligned or connected with gray lines terminating in gray bars). Nucleotides numbered as in (A); those in gray were added to facilitate crystallization. Base-pair geometries indicated using nomenclature of (*27*). (**C**) Ribbon representation of ligase structure, as if peering into the active site (yellow) and ligation junction (red). (**D**) Top-down view, relative to (C).

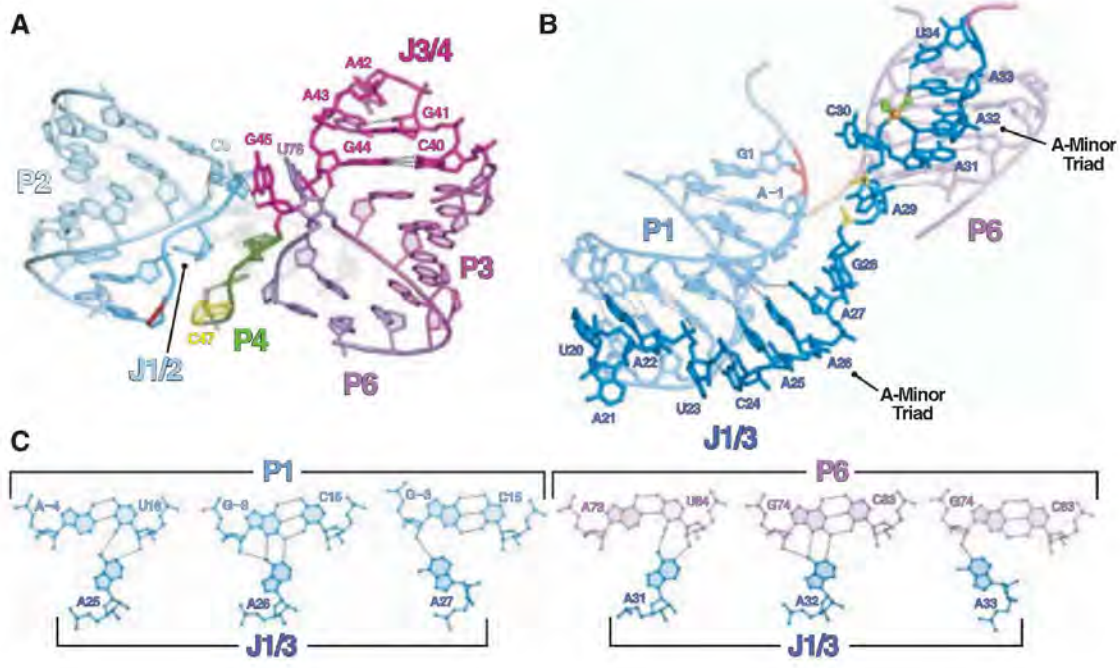


Fig. 2. Tertiary contacts involving the three longest joining regions. (**A**) Interactions bridging the three domains. (**B**) The path of J1/3. (**C**) Hydrogen bonds of the two A-minor triads (fig. S8).

2.2/min in our crystallization conditions (10 mM Mg^{2+}, pH 6.0).

To promote crystallization, we replaced loop 5 (L5) with the U1A-binding loop, and grew crystals of the ligated product complexed with U1A (Fig. 1B) (*11*, *13*). A parallel effort used a phage-display system to generate antibodies for cocrystallization (*14*), which yielded crystals with data to 3.1 Å. The ligase structure in this second crystal form, solved independently, confirmed the structure presented here (all-atom root mean square deviation = 1.48 Å) (fig. S3).

The global structure features three coaxially stacked domains: P1-P2, P3-P6-P7, and P4-P5 (Fig. 1, B and C), consistent with the previously predicted topology (*15*), but with the three domains placed at relative angles of 58° to 71°, converging near the ligation junction so as to resemble a tripod (Fig. 1, C and D). Because the tripod legs protrude into solvent, the fraction of surface area occluded from solvent is less than that of similarly sized RNAs (fig. S4).

Positioning these domains are tertiary interactions at the top of the tripod (Fig. 2A). G45 stacks on U76, the joining residue of the P6-P3 pseudoknot. This interaction pulls the 5′ strands of P4 and P6 close to—and nearly parallel with—the J1/2 joining region, facilitating a contact between a C5 nonbridging oxygen and the 2′-hydroxyl of G45, a group with confirmed function (*10*). Preceding G45 is an unexpected Watson-Crick pair, G44:C40, which we confirmed biochemically (fig. S5) (*11*). This pair extends the P3 helical stack and closes a 3-nucleotide (nt) loop resembling a GNRA tetraloop (in which N is A, C, G, or U; R is A or G). In addition to this loop, two other regions (L7 and part of J1/3) resemble previously defined substructures (figs. S6 and S7).

Interacting with the tripod scaffold is J1/3, which docks into the P1 and P6 minor grooves, passing from one to the other near the ligation junction (Figs. 1C and 2B). Of the 10 minor-groove interactions, eight involve adenosines of J1/3. Residues A25–A26–A27 dock into the fourth and fifth base pairs of P1 (Fig. 2C and fig. S9), which corresponds to the primer-template duplex used by the polymerase. Each interaction could form irrespective of the P1 base-pair identity. Of particular note are the hydrogen bonds involving the 2′-hydroxyls of U16 and G-3. In the polymerase, 2′-deoxy substitution is more detrimental at the position analogous to U16 than at any other template residue, and 2′-deoxy substitution at the position analogous to G-3 is among the most detrimental primer substitutions (*16*). Hence A25–A26–A27 make defined, yet sequence-independent, contacts that help explain the ability of the polymerase to utilize primer-template helices of any sequence (*2*–*4*).

At the other end of J1/3, A31–A32–A33 dock into the P6 minor groove, passing from one helical strand to the other through a succession of hydrogen bonds identical to that of A25–A26–A27 (Fig. 2, B and C). We call this recurring motif the A-minor triad and note another instance in the small subunit of the bacterial ribosome (fig. S8). The P6 A-minor triad helps form a Mg^{2+}-binding site (Figs. 2B and 3A). Direct metal coordination by the A31 and A32 nonbridging (pro-R_P) phosphate oxygens brings these oxygens ~3.1 Å from one another, inducing a 90° kink that positions C30 out of the helical docking register of A31–A33. Outer-sphere contacts involving N7 of A32, N7 and N2 of A33, and O4 of U34 further stabilize this interaction—roles that, in concert with their packing into P6, explain both the absolute conservation of these nucleotides in active ribozyme isolates (*2*, *4*, *8*, *10*, *17*) and the deleterious effects of chemically modifying them (*10*).

Between the two A-minor triads lies the active site (Figs. 2B and 3, A to C). Forming the "floor" of the active site is A71, an absolutely conserved residue at the center of the four-way junction linking the P4-P5 and P3-P6-P7 domains (Fig. 1, A and B). A71 forms an imperfect type I A-minor interaction (*18*) with C86:G105, the first base pair of P7 (Fig. 3A and fig. S9). Chemical modification of A71 or loss of the C86 2′-hydroxyl impairs

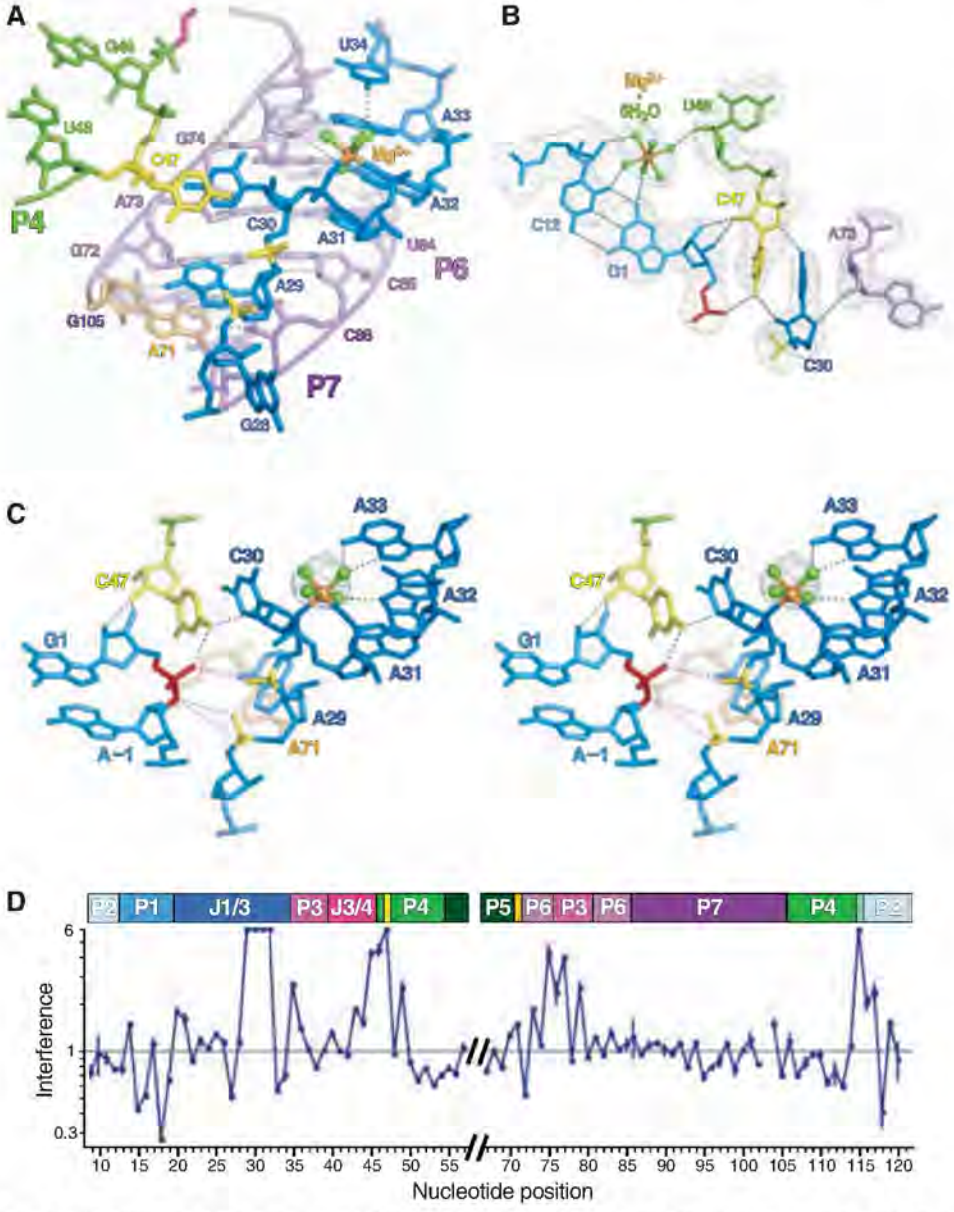


Fig. 3. Architecture of the active site. (**A**) The active site, as viewed from the ligation junction, with P1-P2 removed for clarity. (**B**) Interactions near G1:C12, which is analogous to the NTP-template pair during polymerization (*2*–*4*, *9*). Meshes are simulated-annealing omit maps in which active-site nucleotides (gray, contoured at 2σ) or the hydrated metal cluster (aqua, 4σ) were excluded from map calculations. (**C**) Stereograph of the active site. Black dashes indicate hydrogen bonds; magenta dashes indicate proximity between A29 and C30 phosphate oxygens and the ligation junction (red). Mesh represents a simulated-annealing omit map (4.5 σ) in which the hydrated metal was excluded from map calculations. (**D**) Mean interference values (±SD) from three α-phosphorothiolate NAIM experiments. The secondary structure is aligned above. Interference values were truncated at the detection limit, 6.0 (*10*, *22*). Missing positions are those modified to facilitate crystallization (hashes) or too close to the termini to measure.

catalysis (*10*). A29 stacks on the A71 floor, itself forming an A-minor base triple (fig. S9). Residues A29 and C30 are hence enclosed between the A71 A-minor interaction and the metal-stabilized backbone kink (Fig. 3C). The consequences of this enclosure are twofold. First, C30 is extruded from the minor groove of P6, the rotation of its base constrained by A73 so as to form a cross-strand stack with C47 (Fig. 3, A to C). C47 is likewise extruded from helix P4, the plane of its base roughly perpendicular to the adjoining base pairs, with its N4 exocyclic amine positioned 3.1 Å away from the ligation junction. Second, the A71 interaction prevents G28 from stacking below A29. The consequent rotation of G28 places the phosphates of A29 and C30 in close proximity to one another (~5 Å between phosphorus centers) and facing P1. The nonbridging oxygens at these residues are 4.5 to 5.3 Å from the 3′-hydroxyl and phosphate oxygens of the ligation junction.

We therefore propose that C47, as positioned by C30, and the backbone phosphates of A29 and C30 compose the ligase active site. Both cytidines are conserved among active isolates (*2*, *4*, *8*, *10*, *17*), although their contributions to activity differ (fig. S10). The C30U substitution decreases activity by a factor of five, perhaps from disrupting the hydrogen bond between the C30 N4 and the C47 O4′ (Fig. 3B), whereas the C47U substitution diminishes activity by a factor of $>10^4$, consistent with a more direct role in catalysis.

With only minor perturbation, the A29 and C30 phosphates could provide a binding site for a catalytic metal ion, as observed at active sites for some natural ribozymes (*19*, *20*) and the L1 ligase, an artificial ribozyme that promotes a reaction resembling that of our ligase (*21*). Although we observed no electron density for such a metal ion in the crystal structure of the product, a metal might be bound more tightly before catalysis. To test for a functional role for these and other backbone phosphate oxygens, we performed nucleotide-analog interference mapping (NAIM) (*10*, *22*), randomly incorporating R_P-phosphorothioate substitutions and identifying those that interfered with activity (Fig. 3D). Substitution at residues A29 through A32 resulted in maximal interference. Substitution at A31 and A32 would disrupt the metal-stabilized kink needed to form the C30 to C47 stack at the active site (Fig. 3C). That substitution at A29 or C30 similarly abrogated function supports the hypothesis that these other phosphates coordinate at least one magnesium ion that is catalytically critical in the transition state but not bound tightly in the crystallized product.

We propose a preliminary model for catalysis by the class I ligase and its polymerase derivatives that resembles the mechanism of proteinaceous enzymes. Proteinaceous nucleic acid polymerases require a pair of aspartic acid–bound divalent metal ions supplemented by a general acid that stabilizes and protonates the pyrophosphate leaving group (Fig. 4A) (*23*, *24*). In our model, the substrate α-phosphate and backbone phosphates of A29 and C30 jointly bind a catalytic magnesium ion (Fig. 4B). This metal activates the primer 3′-hydroxyl for nucleophilic attack and stabilizes the transition-state geometry, akin to Metal A of proteinaceous polymerases (*23*). In addition, because free nucleotide triphosphates (NTPs) bind divalent cations, we suggest that the G1 triphosphate (or the incoming NTP) enters the active site complexed with a second metal, which, after binding, would remain coordinated by oxygens on the β- and γ-phosphates. At the transition state, this second metal helps stabilize the developing negative charge on the pyrophosphate leaving group. This stabilization is aided by the exocyclic amine of C47, which hydrogen bonds to the (α,β) bridging oxygen.

Our model, which postulates a hydrogen bond to the leaving oxygen in the transition state, differs from that of proteinaceous polymerases, which involves proton transfer to this oxygen in the transition state (*24*). Although nucleobases can act as general acids at ribozyme active sites (*25*), we disfavor ascribing such a function to C47. With increasing pH, the ribozyme ligation rate increases log-linearly with a slope of 1.0 (pH 5.7 to 8.5), consistent with the net loss of one proton, presumably that of the nucleophile, when proceeding from the ground state to the transition state (*12*). If general-acid catalysis by C47 were dominant at the transition state, the pH–rate profile would likely deviate from linearity over this pH range. Moreover, if C47 were a general acid, the functional group donating the proton would differ from that of the active-site cytidine of the hepatitis delta virus (HDV) ribozyme, wherein the N3 imine is thought to act as the proton donor (*25*). For the ligase, methylating N3 has little effect on catalysis, which rules out direct participation of N3 but not N4 (*10*).

By identifying the residues at the active site, the ligase crystal structure will facilitate directed examination of the catalytic mechanism of RNA-catalyzed RNA polymerization. Our model also provides insights into how known polymerase ribozymes recognize primer-template duplexes, the feature most in need of improvement for developing a self-replicating polymerase ribozyme (*4*, *26*), and one that now can be targeted more explicitly in design and selection experiments.

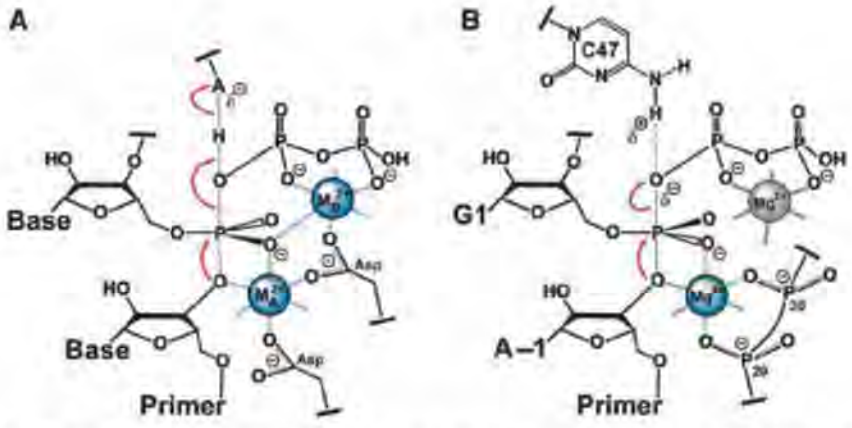


Fig. 4. Transition-state stabilization by polymerases built from either protein or RNA. (**A**) Catalysis by proteinaceous polymerases (*23*, *24*). Indicated are bonds formed or broken during the transition state (red arrows), coordination of catalytic metal ions, M_A and M_B (blue solid lines), and an active-site acid (A•••H). (**B**) Model for catalysis by the ligase ribozyme. Notation as in (A), with the addition of a hydrogen bond between C47 N4 and the leaving group (dashed gray line). Some magnesium ligands are not specified; for those that are, relative orientations are unknown. A proposed contact to the reactive phosphate pro-R_P oxygen (*28*) and two speculative contacts implied by NAIM are in blue. Metal ion and coordinations not supported (or refuted) by structural or biochemical evidence are in gray.

References and Notes

1. G. F. Joyce, L. E. Orgel, in *The RNA World*, R. F. Gesteland, T. R. Cech, J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999), pp. 49–77.
2. W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **292**, 1319 (2001).
3. M. S. Lawrence, D. P. Bartel, *RNA* **11**, 1173 (2005).
4. H. S. Zaher, P. J. Unrau, *RNA* **13**, 1017 (2007).
5. D. P. Bartel, J. W. Szostak, *Science* **261**, 1411 (1993).
6. E. H. Ekland, J. W. Szostak, D. P. Bartel, *Science* **269**, 364 (1995).
7. M. C. Wright, G. F. Joyce, *Science* **276**, 614 (1997).
8. G. F. Joyce, *Annu. Rev. Biochem.* **73**, 791 (2004).
9. E. H. Ekland, D. P. Bartel, *Nature* **382**, 373 (1996).
10. S. C. Bagby, N. H. Bergman, D. M. Shechner, C. C. Yen, D. P. Bartel, *RNA*, published online 27 November 2009 (10.1261/rna.191250).
11. Materials and methods and supporting text are available on *Science* Online.
12. M. E. Glasner, N. H. Bergman, D. P. Bartel, *Biochemistry* **41**, 8103 (2002).
13. A. R. Ferré-D'Amaré, J. A. Doudna, *J. Mol. Biol.* **295**, 541 (2000).
14. J. D. Ye *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 82 (2008).
15. N. H. Bergman, N. C. Lau, V. Lehnert, E. Westhof, D. P. Bartel, *RNA* **10**, 176 (2004).
16. U. F. Müller, D. P. Bartel, *Chem. Biol.* **10**, 799 (2003).
17. E. H. Ekland, D. P. Bartel, *Nucleic Acids Res.* **23**, 3231 (1995).
18. P. Nissen, J. A. Ippolito, N. Ban, P. B. Moore, T. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4899 (2001).
19. M. R. Stahley, S. A. Strobel, *Science* **309**, 1587 (2005).
20. N. Toor, K. S. Keating, S. D. Taylor, A. M. Pyle, *Science* **320**, 77 (2008).
21. M. P. Robertson, W. G. Scott, *Science* **315**, 1549 (2007).
22. S. A. Strobel, *Curr. Opin. Struct. Biol.* **9**, 346 (1999).
23. N. Sträter, W. N. Lipscomb, T. Klabunde, B. Krebs, *Angew. Chem. Int. Ed. Engl.* **35**, 2024 (1996).
24. C. Castro *et al.*, *Nat. Struct. Mol. Biol.* **16**, 212 (2009).

25. P. C. Bevilacqua, R. Yajima, *Curr. Opin. Chem. Biol.* **10**, 455 (2006).
26. M. S. Lawrence, D. P. Bartel, *Biochemistry* **42**, 8748 (2003).
27. N. B. Leontis, E. Westhof, *RNA* **7**, 499 (2001).
28. M. E. Glasner, C. C. Yen, E. H. Ekland, D. P. Bartel, *Biochemistry* **39**, 15556 (2000).
29. We thank A. Ferré-D'Amaré, J. Doudna, R. Batey, and K. Zhou for the generous donation of reagents and protocols; K. Rajashankar for assistance in phasing; E. Duguid, V. Malashkevich, C. Drennan, T. Schwartz, U. RajBhandary, J. Cochrane, M. Stahley, S. Strobel, M. Robertson, and W. Scott for advice, encouragement, and helpful discussions; and members of the Bartel laboratory for support and comments. Supported by NIH grant GM61835. X-ray data were collected with the Northeastern Collaborative Access Team at the Advanced Photo Source, Argonne National Laboratory (*11*). D.M.S. was an NSF predoctoral fellow; S.C.B. was a Howard Hughes Medical Institute (HHMI) predoctoral fellow. D.B. and J.A.P. are HHMI investigators. Coordinates for the U1A-bound and Fab-bound ribozymes are deposited in the Protein Data Bank (3HHN and 3IVK, respectively).

Supporting Online Material
www.sciencemag.org/cgi/content/full/326/5957/1271/DC1
Materials and Methods
SOM Text
Figs. S1 to S12
Tables S1 to S3
References

8 April 2009; accepted 24 July 2009
10.1126/science.1174676

A High-Resolution Structure of the Pre-microRNA Nuclear Export Machinery

Chimari Okada,[1,6]* Eiki Yamashita,[1]* Soo Jae Lee,[2]*† Satoshi Shibata,[3] Jun Katahira,[3,4] Atsushi Nakagawa,[1] Yoshihiro Yoneda,[3,4,5]† Tomitake Tsukihara[1,6]†

Nuclear export of microRNAs (miRNAs) by exportin-5 (Exp-5) is an essential step in miRNA biogenesis. Here, we present the 2.9 angstrom structure of the pre-miRNA nuclear export machinery formed by pre-miRNA complexed with Exp-5 and a guanine triphosphate (GTP)–bound form of the small nuclear guanine triphosphatase (GTPase) Ran (RanGTP). The x-ray structure shows that Exp-5:RanGTP recognizes the 2-nucleotide 3′ overhang structure and the double-stranded stem of the pre-miRNA. Exp-5:RanGTP shields the pre-miRNA stem from degradation in a baseball mitt–like structure where it is held by broadly distributed weak interactions, whereas a tunnel-like structure of Exp-5 interacts strongly with the 2-nucleotide 3′ overhang through hydrogen bonds and ionic interactions. RNA recognition by Exp-5:RanGTP does not depend on RNA sequence, implying that Exp-5:RanGTP can recognize a variety of pre-miRNAs.

Mature microRNAs (miRNAs), short noncoding RNAs present in a wide range of eukaryotes (*1*, *2*), play important roles in the regulation of biological processes including development, cell proliferation, cell differentiation, apoptosis, transposon silencing, and antiviral defense (*3*–*6*). miRNA biogenesis (*7*) begins in the nucleus, where capped and polyadenylated primary miRNAs, several kilobases in length, are transcribed. These are processed by the nuclear ribonuclease (RNase) III enzyme Drosha to generate ~65-nucleotide (nt) pre-miRNAs that have stem-loop structures containing 2-nt 3′ overhangs. Exp-5 translocates pre-miRNAs from the nucleus to the cytoplasm through the nuclear pore complex (*8*–*12*). In the cytoplasm, the pre-miRNAs are further processed by the cytoplasmic RNase III enzyme Dicer, which excises a ~22–base pair (bp) RNA duplex. One strand of the duplex binds to its target mRNA with imperfect complementarity, usually within the target's 3′ untranslated region, assisted by the RNA-induced silencing complex (*7*).

Exp-5 facilitates miRNA biogenesis not only by acting as the nuclear export factor for pre-miRNAs but also by protecting pre-miRNAs from digestion by nucleases. Loss of Exp-5 results in the loss of cytoplasmic miRNA expression without nuclear accumulation of pre-miRNAs (*10*). Pre-miRNA binding to Exp-5 requires the guanine triphosphatase (GTPase) Ran (RanGTP). The Exp-5:RanGTP:pre-miRNA heteroternary complex formed in the nucleus is exported to the cytoplasm. Ran GTPase–activating protein, which promotes guanine triphosphate (GTP) hydrolysis in conjunction with RanBP1 and/or RanBP2, is exclusively localized in the cytoplasm and triggers the conformation change of Ran to induce release of the pre-miRNA cargo from Exp-5 (*13*, *14*).

Here, we report the structure of the Exp-5:RanGTP:pre-miRNA complex at 2.9-Å resolution (Fig. 1A and fig. S1). This complex contains full-length human Exp-5, canine RanGTP residues 1 to 176 (removal of residues 177 to 216 stabilizes the GTP-bound conformation), and the 48-nt human pre-miRNA-30a stem domain, which includes the 2-nt 3′ overhang (nucleotide numbers 1 to 24 and 40 to 63 of human pre-miRNA-30a). Phase information used for the crystal structure analysis was derived from crystals containing Se-methionine–substituted Exp-5, and the RNA sequence was assigned from the Br anomalous signal information in crystals containing pre-miRNA 5-bromo-oxyuracil derivatives. The structure was refined to an *R* factor of 0.247 and free *R* factor of 0.312, and phasing statistics are provided in table S1. We modeled 1082 of 1204 residues of Exp-5. Several loop regions in the 20 HEAT repeats and 55 residues at the C terminus could not be modeled (details in fig. S1), and 13 residues at the C terminus were modeled as a polyalanine α helix. The residues 1 to 6 of Ran were not modeled because of their disordered structure. Electron density for the pre-miRNA was detected for nucleotides 1 to 11, 14 to 24, and 40 to 63 (fig. S2). The pre-miRNA-30a adopted a typical A-form RNA helical structure, 60 Å in length and 20 Å in diameter.

The Exp-5:RanGTP:pre-miRNA complex is an ellipsoid with dimensions of 65 Å by 80 Å by 110 Å. The crystal structure contains two ternary complexes, labeled A and B, in the asymmetric unit, which are essentially similar [root mean square (RMS) of 1.84 Å, where B is slightly more open than A] and present the same recognition modes for the pre-miRNA. Detailed structural comparison of ternary complexes A and B is described in (*15*). The structure of Exp-5 resembles a tightly wound spring, as seen in other members of the importin-β family. Such conformations are expected to be intrinsically flexible, so small changes in the relative orientation of successive HEAT repeats could cumulatively generate substantial changes in the helicoidal pitch (*16*). Ternary complex A yielded more contrast in its electron density map than did complex B; thus, all structural descriptions of the ternary complex in the following discussion will be restricted to ternary complex A. The Exp-5:RanGTP complex forms a baseball mitt–like structure in which the pre-miRNA is packed (Fig. 1B). A tunnel-like structure at the bottom of the mitt connects the inner space of the mitt with the outer space (Fig. 1B).

The pre-miRNA stem is caught in the mitt formed by the Exp-5:RanGTP complex (Fig. 1), whereas the 15-Å 2-nt 3′ overhang is inserted into a tunnel formed from elements of HEAT repeats 12 to 15 (Figs. 2 and 3 and fig. S3). The inner surface of the tunnel is positively charged

[1]Institute for Protein Research, Osaka University, 3-2 Yamada-oka, Suita, Osaka 565-0871, Japan. [2]College of Pharmacy, Chungbuk National University, Gaeshin-dong, Heungduk-gu, Cheongju City, Chungbuk, Korea. [3]Department of Frontier Biosciences, Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan. [4]Department of Biochemistry, Graduate School of Medicine, Osaka University, 2-2 Yamada-oka, Suita, Osaka 565-0871, Japan. [5]Japan Science and Technology Agency (JST), Core Research for Evolutional Science and Technology (CREST), Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan. [6]Department of Life Science, University of Hyogo, 3-2-1 Koto, Kamighori, Akoh, Hyogo 678-1297, Japan.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: sjlee@chungbuk.ac.kr (S.J.L.); yyoneda@anat3.med.osaka-u.ac.jp (Y.Y.); tsuki@protein.osaka-u.ac.jp (T.T.)

(Figs. 1B and 3B) and probably stabilizes the negatively charged 2-nt 3′ overhang structure. At the tunnel entrance, located at the bottom of the mitt, the guanidyl group of Arg^{602} (HEAT12), which is engaged in tight π-π stacking with the base pair of G1:C61 (Fig. 3 and fig. S3), sterically blocks the double-stranded stem from entering the tunnel. The 2-nt 3′ overhang structure in the tunnel is stabilized by a number of hydrogen bonds and salt bridges with amino acid residues of Exp-5, as shown in table S2, Fig. 3, and fig. S3. Because all interactions involve atoms of the sugar-phosphate backbone, 2-nt 3′ overhang recognition by Exp-5 is independent of RNA sequence.

A previous study of the 3′ overhang structure of pre-miRNA-30a showed that mutants containing 3′ overhangs with a length between 1 and 8 nt had similar binding activity to Exp-5 (*17*, *18*). It is probably because the three additional nucleotides of the 5-nt 3′ end would pass through the aperture to maintain the interaction between Exp-5 and the 2-nt 3′ overhang of the pre-miRNA (Figs. 1A and 3).

In contrast to the above observations concerning the 3′ end of the RNA, in vitro Exp-5 binding assay showed that mutants containing a 5′ overhang of either 2 or 5 nt have reduced binding affinity to Exp-5 (*17*). A model of such a mutant with a 5′ overhang built with the crystal structure as a template suggests that the 3′ end of the stem would sterically clash with HEAT 11 and 12 if the 5′ overhang were to be inserted into the tunnel (Fig. 4). The modeling thus suggests how Exp-5 might discriminate between 3′ and 5′ end structures.

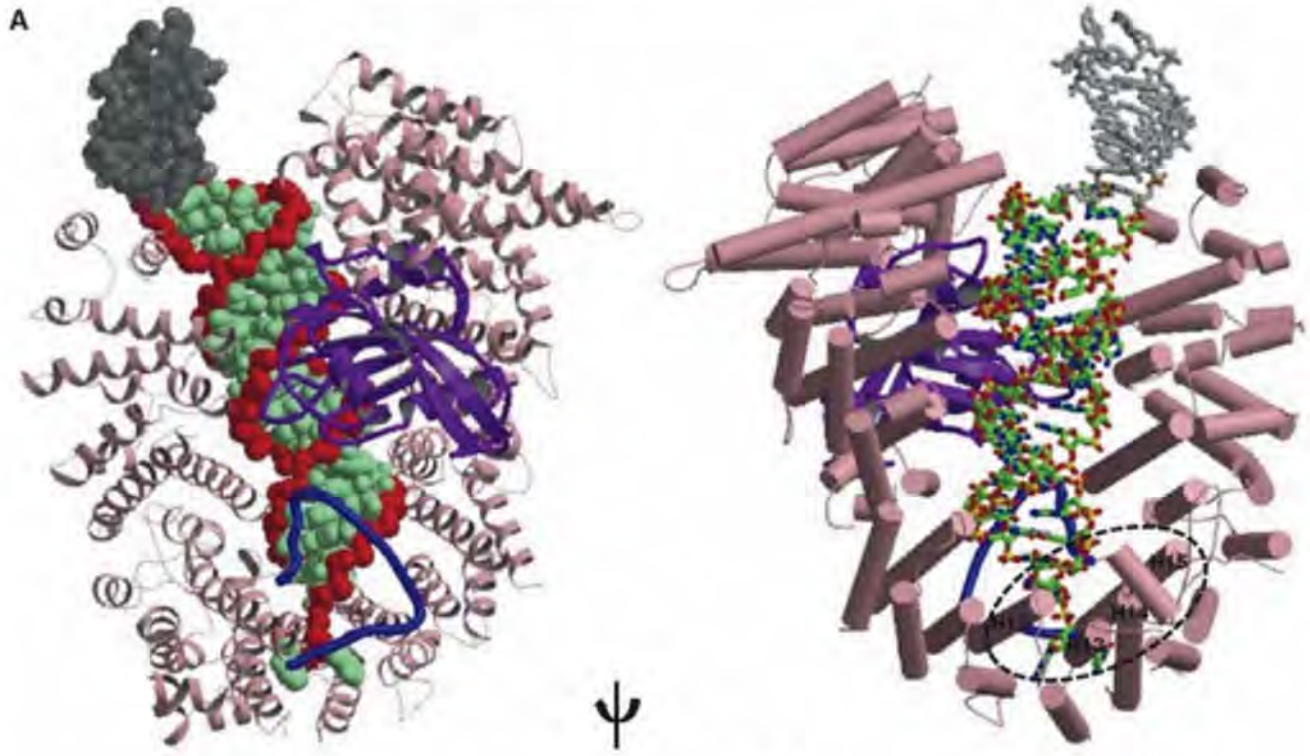


Fig. 1. The structure of the Exp-5:RanGTP:pre-miRNA-30a complex. (**A**) (Left) The structure shows human pre-miRNA-30a (red and green) bound to Exp-5 (pink) with RanGTP (purple). The noncrystalline portion of the pre-miRNA is represented in gray. The long loop of HEAT15 is shown as a blue wire. At the front of this view, the pre-miRNA is shielded by the long loop of Exp-5 and RanGTP. (Right) The pre-miRNA molecule is viewed from the front open side. This is the standard view presented in this paper. Exp-5 covers most of the stem moiety, and the 2-nt 3′ overhang structure (circled black) has many interactions with HEAT repeats 12 to 15, each of which is given a letter "H" with a corresponding number. (**B**) Electrostatic surface potentials of Exp-5:RanGTP calculated by using eF-site (*26*). The potentials are represented in a color gradient from red to blue for the vertex with the potential from −0.1 V to 0.1 V. The black backbone represents the stem moiety of the pre-miRNA to facilitate the display of the pre-miRNA binding site. The inner surface of the mitt is positively charged, and no intensively acidic area is found on the inner surface of Exp-5, distinguishing it from other importin-β family proteins.

Although Exp-5 is an acidic protein with isoelectric point (pI) = 5.6, the protein has localized positive charges from 27 basic residues on the inner surface of the mitt that interact with the negatively charged double-stranded RNA (dsRNA) (Fig. 1B). This contrasts with the related importin-β, which has a negatively charged inner surface to accept the basic IBB domain of importin-α (*19*) (fig. S4). Of the 27 basic residues on the inner surface of the mitt, 11 arginine residues, 1 lysine residue, and 1 histidine residue are involved in 21 interactions with the 16-bp stem of the pre-miRNA molecule within an interatomic distance of 4 Å, as listed in table S3. The interacting residues of Exp-5 are distributed broadly on the inner helices of HEAT 6 to 19 and a loop of HEAT15 to recognize the outer phosphodiester group of the pre-miRNA 16-bp stem (Fig. 2 and fig. S5). By interacting with phosphate groups of the double-stranded pre-miRNA, Lys^{132} and Lys^{134} of RanGTP bridge phosphate groups of a major groove. The hydrogen bond between Glu^{445} carboxyl group and guanine base of G55 is one of two hydrogen bonds between amino acid side chains and bases in the double-stranded stem of the pre-miRNA. It does not appear to be base-specific, as judged by the hydrogen bond geometry. The dipole moment of the inner helix of HEAT19 is attracted to the negative charge of the backbone (fig. S5), as occurs in *Xenopus laevis* RNA-binding protein A (*20*). The interatomic distances between the RNA stem and the proteins are significantly longer than those within the tunnel region, and the number of contacts per nucleotide in the mitt is notably fewer than those in the tunnel (table S3 and fig. S5). Thus, the 16-bp stem of the pre-miRNA, 45 Å in length, is most likely roughly recognized through a broad range of positively charged inner surface residues of the Exp-5:RanGTP mitt.

This structural study indicates that pre-miRNA interacts with Exp-5:RanGTP mainly through ionic interactions in an RNA sequence–independent manner. This is consistent with an experiment that shows that a high ionic strength medium [containing 50 mM $MgCl_2$ at pH = 7.4, 20 mM tris-HCl, and 1 mM dithiothreitol (DTT)] releases pre-miRNA from the Exp-5:RanGTP:Pre-miRNA ternary complex. RNA sequence–independent recognition by Exp-5:RanGTP is supported by in vitro binding assays, which demonstrated that the binding of pre-miRNA to Exp-5 occurs through the double-stranded stem and protruded 2-nt 3′ overhang but does so independently of RNA sequence (*17*).

Because both 3′ and 5′ ends of the pre-miRNA are completely shielded in the tunnel, the pre-

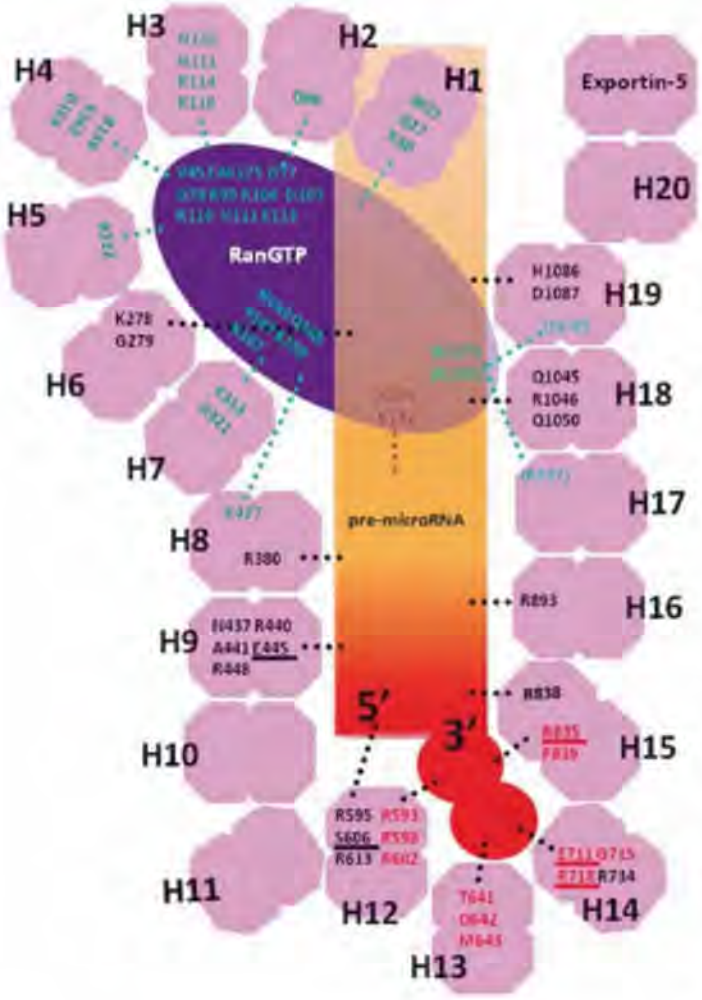


Fig. 2. Intermolecular interactions in the Exp-5:RanGTP:pre-miRNA-30a complex are shown schematically. The pre-miRNA is colored in red and light brown, where red regions interact with more basic residues (*27*) of Exp-5 than do light brown regions. A pink double octagon represents each HEAT repeat. Underlined residues interact with bases of RNA with interatomic distance shorter than 3.5 Å. Exp-5 recognizes pre-miRNA through charged residues such as Arg and Lys. Exp-5 residues shown in red are involved in the recognition of the 2-nt 3′ end of the pre-miRNA, and black residues indicate interactions with the double-stranded stem of the pre-miRNA. Green residues indicate interactions between Exp-5 and RanGTP.

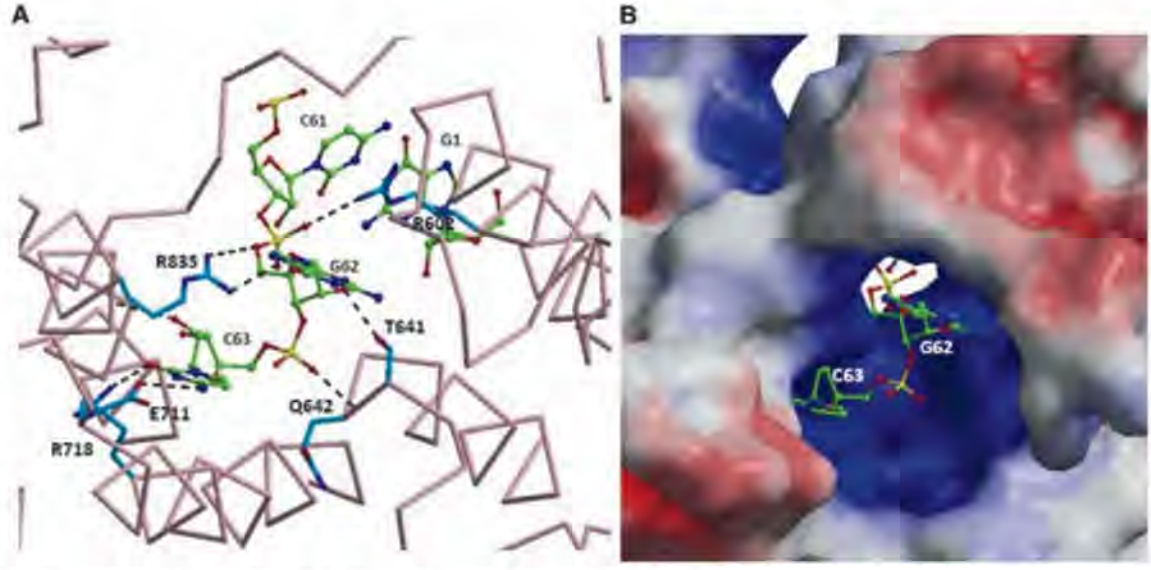


Fig. 3. The structure of 2-nt 3′ overhang of the pre-miRNA in the tunnel viewed from outside of the Exp-5 molecule. (**A**) Hydrogen bonds or salt bridges in the tunnel are represented by broken lines. Exp-5 is colored in pink for Cα, deep blue for N, red for O, and sky blue for the other side chain atoms. RNA is colored in deep blue for N, red for O, yellow for P, and green for the other atoms. Exp-5 residues [Arg^{602}, Thr^{641}, Gln^{642}, Arg^{718}, and Arg^{835} (*27*)] have hydrogen bonds or salt bridges to pre-miRNA nucleotides (G62 and C63) with interatomic distances shorter than 3.5 Å. (**B**) Electrostatic potential is represented as in Fig. 1B. Inside of the tunnel is electrostatically basic and shows many hydrophilic interactions with the 2-nt 3′ overhang structure.

miRNA is protected from digestion by exonucleases.The structure of Exp-5:RanGTP:pre-miRNA indicates that Exp-5:RanGTP prevents an endonuclease from approaching the pre-miRNA in the ternary complex (Fig. 1 and fig. S6). Exp-5:RanGTP surrounds the pre-miRNA on four sides, protecting it from ribonuclease digestion during export from the nucleus to the cytoplasm. Consistent with this, cytoplasmic miRNA expression is controlled by Exp-5 expression, and low Exp-5 expression leads to nuclear degradation of pre-miRNAs (*10*). Thus, Exp-5:RanGTP may act as both a nuclear export carrier and a molecular stabilizer for pre-miRNAs.

It is known that Exp-5 exports not only pre-miRNAs but also other small structured RNAs, such as tRNAs, human Y1 RNA, and adenovirus VA1 RNA, all of which have 3′ overhang structures (*18*, *21*). Formation of the corresponding Exp-5:RanGTP:RNA complexes was examined by building structural models of the ternary complexes for these RNAs with use of the ternary complex of Exp-5:RanGTP:pre-miRNA as a template. These small RNAs could be efficiently packed into both the basic inner surface of the mitt and the tunnel. In the resulting models, the tunnel exhibiting highest positive charge of the molecule recognizes a 3′ overhang structure carrying negative charge; the inner basic surface of the mitt surrounds the sugar-phosphate backbone of double-stranded stem moieties; and lastly, the mitt accommodates small protrusions at the center of the stem in the mitt by low-energy conformational changes of Exp-5 created by a springlike movement. Detailed structure predictions are described in the SOM. A structure of Exportin (Xpot) from *Schizosaccharomyces pombe* in complex with RanGTP and tRNA was recently reported by Cook *et al.* (*22*). No sequence homology was detected between Exp-5 and Xpot, and Xpot does not exhibit a structure corresponding to the tunnel of Exp-5 that recognizes the 3′ overhang structure. In addition, the orientation of the RNA stem in the Xpot:RanGTP complex is inverse to that in Exp-5:RanGTP.

In radiolabel binding studies of Exp-5:RanGTP, five times more human Y1 RNA than VA1 RNA was required to compete for complex formation (*18*). This may be due to the fact that VA1 has only four single-nucleotide protrusions, so the larger protrusion in human Y1 may explain the inefficiency of the Exp-5:RanGTP:RNA complex formation compared with that of VA1 resulting from steric hindrance.

The truncated RanGTP superimposed well on the corresponding region of full-length RanGTP [Protein Data Bank (PDB) under accession code 1IBR] with an RMS deviation of 0.44 Å for the Cα atoms. Exp-5 interacts with RanGTP at three sites, and the detailed interaction is shown in Fig. 2. The first Ran binding site of Exp-5 is formed by HEAT repeats 1 to 5 and wraps around the Ran switch II loop similarly to other importin-β family molecules (*23*, *24*). The second Ran-binding interface of Exp-5 is formed by HEAT repeats 7 and 8 and is similar to the binding interface of a yeast importin-β (Kap95p) and CRM1 (*25*). The third Ran-binding interface of Exp-5 is formed by HEAT repeats 17 and 19, which is different from those of other published importin-β family proteins (*23*–*25*). The total buried surface between Exp-5 and RanGTP (1725 Å^2) is the smallest of all previously reported RanGTP-binding complexes: The buried surface in the Cse1:RanGTP:Kap60p structure was 1822 Å^2, that of yeast importin-β Kap95p:RanGTP was 2044 Å^2, and that of CRM1:RanGTP:Snurportin1 was 2227 Å^2. The smallest buried surface of Exp-5 among the importin-β family is primarily due to the difference in the third interaction site. The few contacts between RanGTP and HEAT repeats 17 and 19 might contribute to the relatively high flexibility of the mitt structure.

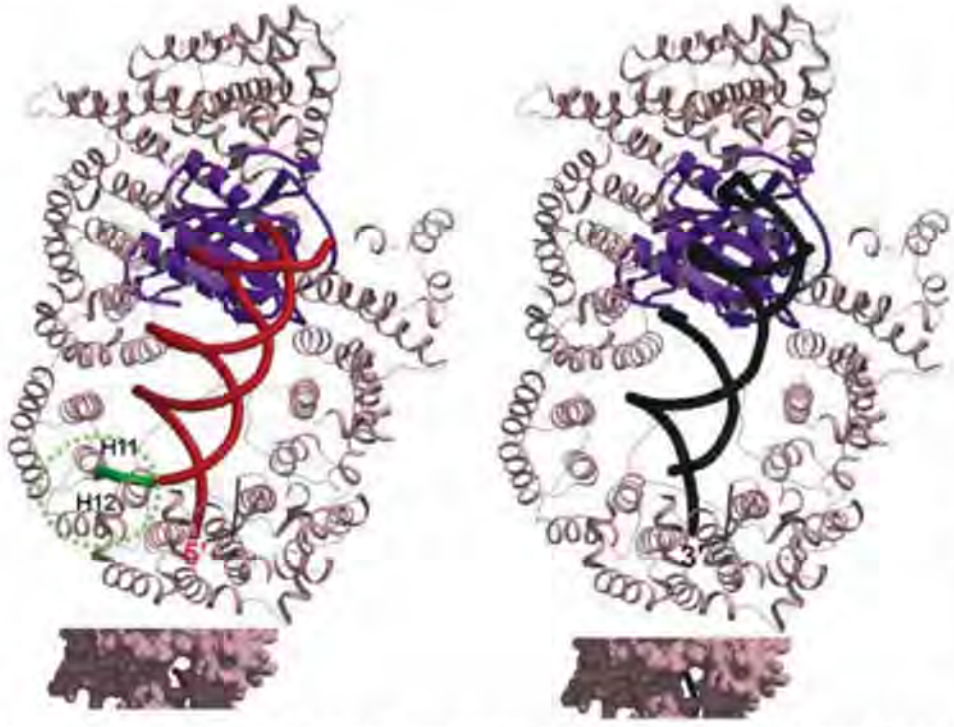


Fig. 4. The predicted model of a 2-nt 5′ overhang double-stranded RNA (top left) is depicted in the same orientation as Exp-5:RanGTP:pre-miRNA (top right). The model of 2-nt 5′ overhang double-stranded RNA is colored in green for its 5′ overhang region and red for the other part. Their 5′ end and 3′ end in the tunnel are depicted in surface representations viewed from outside of the mitt (bottom left and bottom right, respectively). Surface was calculated by a program MSMS (*28*). The 2-nt 5′ overhang is inserted into and the 3′ end of the stem sterically clashes with HEAT 11 and 12 in the predicted structure. The clashed regions are highlighted by the green dotted circle.

References and Notes

1. R. C. Lee, R. L. Feinbaum, V. Ambros, *Cell* **75**, 843 (1993).
2. B. Wightman, I. Ha, G. Ruvkun, *Cell* **75**, 855 (1993).
3. V. Ambros, *Nature* **431**, 350 (2004).
4. G. B. Ruvkun, *Harvey Lect.* **99**, 1 (2003).
5. B. R. Cullen, *Nature* **457**, 421 (2009).
6. A. Ventura, T. Jacks, *Cell* **136**, 586 (2009).
7. V. N. Kim, J. Han, M. C. Siomi, *Nat. Rev. Mol. Cell Biol.* **10**, 126 (2009).
8. A. M. Brownawell, I. G. Macara, *J. Cell Biol.* **156**, 53 (2002).
9. E. Lund, S. Güttinger, A. Calado, J. E. Dahlberg, U. Kutay, *Science* **303**, 95 (2004); published online 20 November 2003 (10.1126/science.1090599).
10. R. Yi, Y. Qin, I. G. Macara, B. R. Cullen, *Genes Dev.* **17**, 3011 (2003).
11. M. T. Bohnsack, K. Czaplinski, D. Gorlich, *RNA* **10**, 185 (2004).
12. S. Shibata *et al.*, *Nucleic Acids Res.* **34**, 4711 (2006).
13. B. R. Cullen, *J. Cell Sci.* **116**, 587 (2003).
14. V. N. Kim, *Trends Cell Biol.* **14**, 156 (2004).
15. Materials and methods are available as supporting material on *Science* Online.
16. E. Conti, C. W. Muller, M. Stewart, *Curr. Opin. Struct. Biol.* **16**, 237 (2006).
17. Y. Zeng, B. R. Cullen, *Nucleic Acids Res.* **32**, 4776 (2004).
18. C. Gwizdek *et al.*, *J. Biol. Chem.* **278**, 5505 (2003).
19. G. Cingolani, C. Petosa, K. Weis, C. W. Muller, *Nature* **399**, 221 (1999).
20. J. M. Ryter, S. C. Schultz, *EMBO J.* **17**, 7505 (1998).
21. A. Calado, N. Treichel, E. C. Muller, A. Otto, U. Kutay, *EMBO J.* **21**, 6216 (2002).
22. A. G. Cook, N. Fukuhara, M. Jinek, E. Conti, *Nature* **461**, 60 (2009).
23. Y. Matsuura, M. Stewart, *Nature* **432**, 872 (2004).
24. S. J. Lee, Y. Matsuura, S. M. Liu, M. Stewart, *Nature* **435**, 693 (2005).
25. T. Monecke *et al.*, *Science* **324**, 1087 (2009); published online 23 April 2009 (10.1126/science.1173388).
26. K. Kinoshita, H. Nakamura, *Bioinformatics* **20**, 1329 (2004).
27. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
28. The URL of the MSMS program is www.scripps.edu/~sanner/html/msms_home.html.

29. This work was supported in part by grants-in-aid for scientific research (16087101, 16087206, and 21227003), the GCOE program (A-041) from the Ministry of

Education, Culture, Sports, Sciences, and Technology of Japan (to T.T.), the Strategic Japan-UK Cooperation Program of the Japan Science and Technology Agency (to T.T.), a grant from Takeda Science Foundation (to Y.Y.), and a grant from Chungbuk BIT Research-Oriented University Consortium (to S.J.L.). Coordinates and structure factures have been deposited to the PDB under accession code 3A6P.

Supporting Online Material
www.sciencemag.org/cgi/content/full/326/5957/1275/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S3
References

6 July 2009; accepted 9 October 2009
10.1126/science.1178705

Crystal Structure of a Nucleocapsid-Like Nucleoprotein-RNA Complex of Respiratory Syncytial Virus

Rajiv G. Tawar,[1]* Stéphane Duquerroy,[1,2]* Clemens Vonrhein,[3] Paloma F. Varela,[1]† Laurence Damier-Piolle,[1] Nathalie Castagné,[4] Kirsty MacLellan,[5]‡ Hugues Bedouelle,[6] Gérard Bricogne,[3] David Bhella,[5] Jean-François Eléouët,[4]§ Félix A. Rey[1]§

The respiratory syncytial virus (RSV) is an important human pathogen, yet neither a vaccine nor effective therapies are available to treat infection. To help elucidate the replication mechanism of this RNA virus, we determined the three-dimensional (3D) crystal structure at 3.3 Å resolution of a decameric, annular ribonucleoprotein complex of the RSV nucleoprotein (N) bound to RNA. This complex mimics one turn of the viral helical nucleocapsid complex, which serves as template for viral RNA synthesis. The RNA wraps around the protein ring, with seven nucleotides contacting each N subunit, alternating rows of four and three stacked bases that are exposed and buried within a protein groove, respectively. Combined with electron microscopy data, this structure provides a detailed model for the RSV nucleocapsid, in which the bases are accessible for readout by the viral polymerase. Furthermore, the nucleoprotein structure highlights possible key sites for drug targeting.

Human respiratory syncytial virus (RSV) is an important viral agent of pediatric respiratory tract disease worldwide, causing pneumonia and bronchiolitis in infants (*1*). No vaccine is currently available, and effective treatments have yet to be developed (*2*). RSV is a nonsegmented, negative-strand RNA virus of the *Paramyxoviridae* family in the *Mononegavirales* order (*3*), which also includes the *Rhabdoviridae*, *Bornaviridae*, and *Filoviridae* families. RSV is further classified into the *Pneumovirus* genus within the *Pneumovirinae* subfamily of the *Paramyxoviridae* (*4*). The 15.2-kb genomic RNA contains 10 genes, 4 of which code for intracellular proteins that are involved in genome transcription, replication, and particle budding and have orthologs in all *Mononegavirales*, namely N (nucleoprotein), P (phosphoprotein), M (matrix protein), and L ("large" protein, containing the RNA polymerase catalytic motifs).

The RSV genomic RNA forms a nuclease-resistant (*1*) helical ribonucleoprotein (RNP) complex with the N protein, termed nucleocapsid, which is used as template for RNA synthesis by the viral polymerase complex (*5*). Studies on Sendai virus [a paramyxovirus (*6*)] and on the vesicular stomatitis virus [VSV, *Rhabdoviridae* family (*7*)] have shown that the polymerase in replication mode includes an L-P-N core complex, whereas in transcription mode, it is a complex of L, P, and cellular proteins. This pattern is believed to hold true for all members of the *Mononegavirales*. Electron microscopy (EM) observations of authentic RSV nucleocapsids derived from infectious virions (*8*), as well as recombinant nucleocapsid-like RNP complexes (*9*), revealed flexible helical filaments 15 nm in diameter, with spikes protruding at 45°, resembling a head of wheat.

Structural data are necessary to understand the mechanism of RNA synthesis used by these viruses. The structure of L (>200 kD) is not known for any of the *Mononegavirales*. Three-dimensional (3D) structures of fragments of P are available for a number of them, but not for RSV. The crystal structure of N was determined for the Borna disease virus [BDV, *Bornaviridae* family (*10*)] and for the rhabdoviruses VSV (*11*) and rabies virus [RV (*12*)]. The latter two structures revealed decamer and hendecamer N rings, respectively, which were bound nonspecifically to cellular RNA. Extrapolation from these structures to understand how the RNA in the nucleocapsid is presented to the VSV and RV polymerases for viral RNA synthesis was not straightforward, however, because the RNA was occluded inside the rings.

We obtained crystals of recombinant decameric RSV N-RNA rings diffracting to 3.3 Å resolution (*13*). Phase extension from a low-resolution EM reconstruction (*14*) using 20-fold real-space averaging yielded a very clear 3.3 Å–resolution electron density map in which the atomic model was built and refined (table S1). Despite the relatively limited resolution, the 20-fold redundancy made the positioning of atoms sufficiently accurate for an unambiguous interpretation of their interactions (hydrogen bonds, for instance), which would not be the case in a less-over-determined structure at 3.3 Å resolution.

The RNA chain runs within a basic surface groove, surrounding the periphery of the N protein ring, as a belt (Fig. 1 and fig. S2). Each N subunit interacts with seven ribonucleotides, resulting in an RNA chain of 70 bases in total. The RNA electron density is well defined, with the density of the bases corresponding to the mean of the four bases of the genetic code, reflecting the random composition of the bound RNA (*15*). We modeled the bases as citosines. The RNA adopts an extended conformation with phosphate-sugar torsion angles and sugar pucker similar to those of A-DNA (table S2), except at two "switches."

Each N subunit is organized as a core region containing two domains, N- and C-terminal (NTD and CTD), that are connected through a hinge region. The RNA groove is formed at the NTD/CTD interface, the interdomain connection lines its internal side, and the α_{C3}-to-α_{C4} loop (Fig. 1) forms a lid partially covering the exposed side of the RNA. This NTD/CTD core of the molecule has N- and C-terminal extensions, termed N arm and C arm, respectively, that appear folded only in the context of the quaternary interactions in the ring. A search of the Protein Data Bank with the Dali server (*16*) indicated that the closest relatives are, as expected, the BDV, VSV, and RV N protein orthologs (table S3 and Fig. 2). Both the NTD and CTD interact laterally in the ring with their counterparts from the adjacent subunits but do not make tight contacts. The interacting surfaces are highly hydrated, primarily involving intermittent van der Waals contacts. The ring is stabilized by the RNA belt and by the N chain that results from the insertion of the N arm (residues 1 to 35) of one subunit into the

[1]Institut Pasteur, Unité de Virologie Structurale, Département de Virologie and CNRS Unité de Recherche Associée (URA) 3015, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France. [2]Université Paris-Sud, Faculté d'Orsay, 91405 Orsay Cedex, France. [3]Global Phasing, Sheraton House, Castle Park, Cambridge CB3 0AX, UK. [4]L'Institut National de la Recherche Agronomique, Unité de Virologie et Immunologie Moléculaires, Domaine du Vilvert, 78350 Jouy-en-Josas, France. [5]Medical Research Council Virology Unit, University of Glasgow, Church Street, Glasgow G11 5JR, UK. [6]Institut Pasteur, CNRS URA 3012, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France.

*These authors contributed equally to this work.
†Present address: Synchrotron Soleil, Laboratoire de Biologie, L'Orme des Merisiers, Saint Aubin, Boite Postale 48, 91192 Gif-sur-Yvette Cedex, France.
‡Present address: National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG, UK.
§To whom correspondence should be addressed. E-mail: jean-françois.eleouet@jouy.inra.fr (J.-F.E.); rey@pasteur.fr (F.A.R.)

compact fold of the adjacent one (Fig. 1B). This lateral connectivity helps explain the observed malleability of the N-N interactions in the flexible, yet very stable, RSV nucleocapsid. The C arm lies above the CTD in the ring, occupying the space that would be between consecutive turns of the helical nucleocapsid. It contains residues 361 to 391, the last 12 to 20 amino acids being disordered in the crystal (*13*).

Both core domains are α-helical bundles, with 10 α helices in the NTD and 4 in the CTD. The NTD has 218 residues (36 to 253), with a long β hairpin projecting away from the molecule at the most distal end (Fig. 2). This region is not conserved in amino acid sequence within the *Pneumovirinae* (fig. S3) and is also most variable in 3D structure when compared with the other mononegavirus N proteins (Fig. 2). Mutations conferring in vitro resistance to an anti-RSV compound, RSV604 (*17*), map to this insertion in strand β_{I2} and in helix α_{I2} (Fig. 2 and fig. S3).

The CTD has 107 residues (254 to 360), spanning the most conserved region in primary structure within the *Pneumovirinae*. Furthermore, the structural alignment shows that its four α helices pack identically in the available N structures of the various mononegaviruses, in spite of the lack of sequence conservation. The α_{C3}-α_{C4} loop (Fig. 1C) is responsible for many of the RNA contacts, with the side chain of Arg^{338} arching above the RNA to make a salt bridge with Asp^{175} in the NTD (fig. S2B).

The 7 nucleotides (nt) interacting with each N subunit have bases 2 to 4 stacked and facing the protein in a cavity within the groove, at the N-N interface (Fig. 1 and fig. S2C). Base 1 is sandwiched between the upstream base 7 and the backbone of helix α_{N8} (Fig. 1C). This α helix has two glycine residues in consecutive turns (Gly^{241} and Gly^{245}), which align to form a flat face of the helix on which base 1 packs. The packing of base 1 on the preceding base 7 results in a row of four stacked bases (5-6-7-1) facing solvent, away from the protein. The 2-3-4 base stack in the cavity is such that base 2 contacts Asn^{249} and Arg^{185}, and base 4, on the other side, contacts Trp^{260} and Val^{256}. Base 3, which is in the middle, makes no direct protein contacts. The RNA conformation switches from base out to base in at phosphate 1, and vice versa at phosphate 4 (going 3′ to 5′), constrained by the presence of helix α_{N8} at switch 1, and that of Tyr^{337}, which packs against the ribose ring of nucleotide 4, at switch 2 (Fig. 1C). Mutation of Tyr^{337} was reported to abolish nucleocapsid formation (*18*), highlighting the key role of its aromatic side chain in imposing the required conformation on the RNA chain. As indicated in Fig. 1D, all the observed hydrogen bonds are directed to the RNA backbone. In particular, the 2′ OH of riboses 4, 5, and 6 donate hydrogen bonds to main-chain carbonyls, explaining the specificity for RNA instead of DNA.

The 3D superposition of the VSV N subunit onto its RSV counterpart reveals that the three-base stack of RNA facing the protein binds in the same way in the two complexes in a cavity at the N-N interface, in spite of the presence of a bulged-out base in the VSV complex (Fig. 3C). Although the RNA groove is in the same location in both cases, the lateral N contacts in the VSV and RV rings are such that the curvature is opposite to that of the RSV ring (Fig. 3). The result is an inside-out nucleocapsid ring, with the RNA inside and the N molecule oriented outside-in. Because each rhabdovirus N subunit contacts nine ribonucleotides instead of seven,

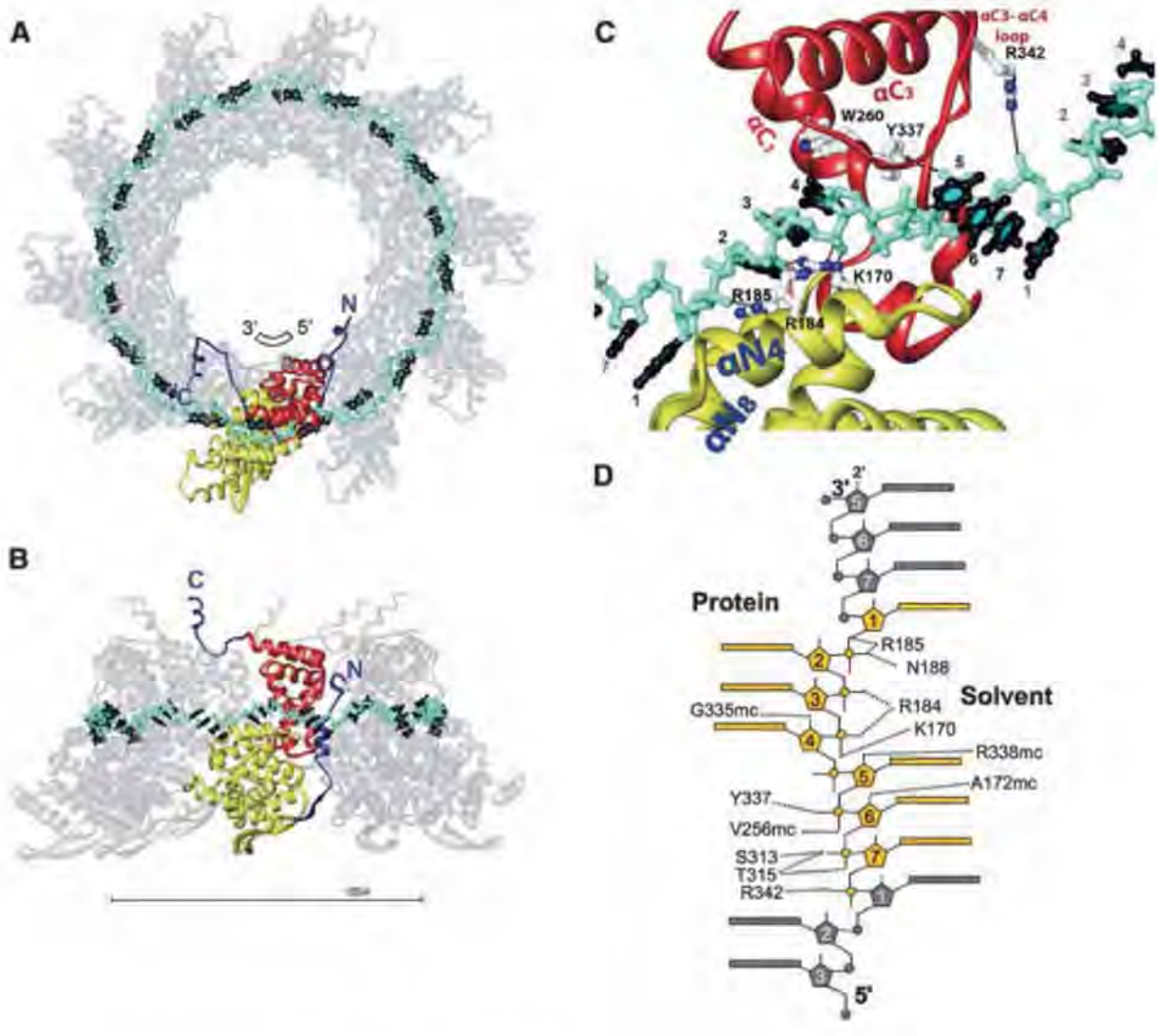


Fig. 1. A decameric ribonucleoprotein ring complex. One N subunit is colored according to domains: yellow and red indicate NTD and CTD, respectively, and the N and C arms are blue. The RNA is displayed with the backbone in cyan and the bases in black, except in (D). (**A**) View down the ring axis. The RNA polarity is indicated within the ring. (**B**) Side view. The bar indicates 100 Å. (**C**) Close-up of the protein-RNA interactions. (**D**) Details of the interactions with RNA. The schematic diagram is numbered from 3′ to 5′. Pentagons, rectangles, and small circles represent riboses, bases, and phosphates, respectively. The relevant contacts with the protein are indicated. Main-chain contacts are indicated by mc.

the RNA ring follows a more convoluted path to fit 90 nt in a smaller diameter than the 70-nt ring of the RSV counterpart (Fig. 3).

Although the N RNA contacts in the groove are not base-specific, the cavity appears tailored to bind a set of three stacked bases, a feature that appears to be conserved across the *Mononegavirales* order. Because the bases are averaged out in our crystals, it is not possible to tell from the structure whether certain particular nucleotide sequences would make stronger or weaker interactions within the cavity. The overall arrangement of the RNA around the ring is reminiscent of the structure of the *trp* RNA-binding attenuation protein (TRAP)/RNA complex from *Bacillus subtilis* (*19*), which also forms a ring of 11 TRAP subunits, with the RNA running at the periphery, and three stacked bases are inserted into a cavity located at the subunit interface. However, in contrast to the TRAP/RNA complex, the contacts with RSV N are not base-specific but rather RNA backbone–specific.

We calculated a 26 Å–resolution EM reconstruction (*13*) from cryo-negative stain images

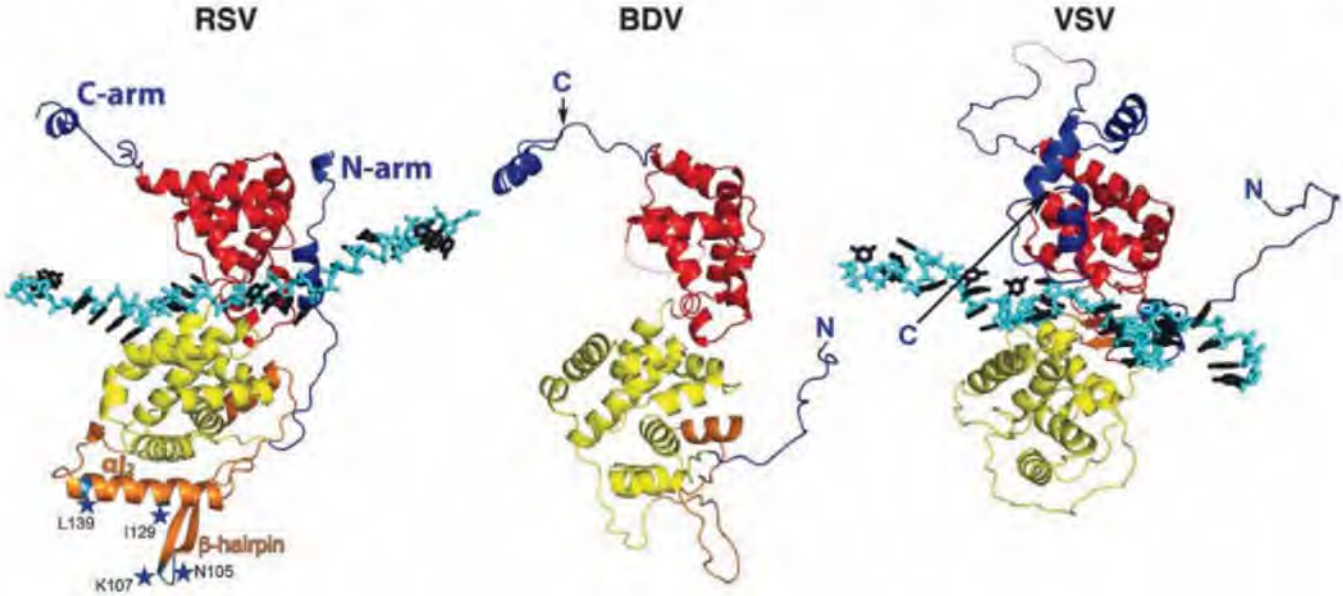


Fig. 2. 3D fold of RSV N and comparison with BDV and VSV N. (**Left**) RSV N is colored according to domains as in Fig. 1. The variable region is in orange, with the β hairpin highlighted and labeled with the RSV604 resistance mutation sites. The BDV (**center**) and VSV (**right**) N structures, respectively, oriented and colored identically to RSV N. The conserved domains (used for the alignment) are red and yellow. The RNA is colored as in Fig. 1.

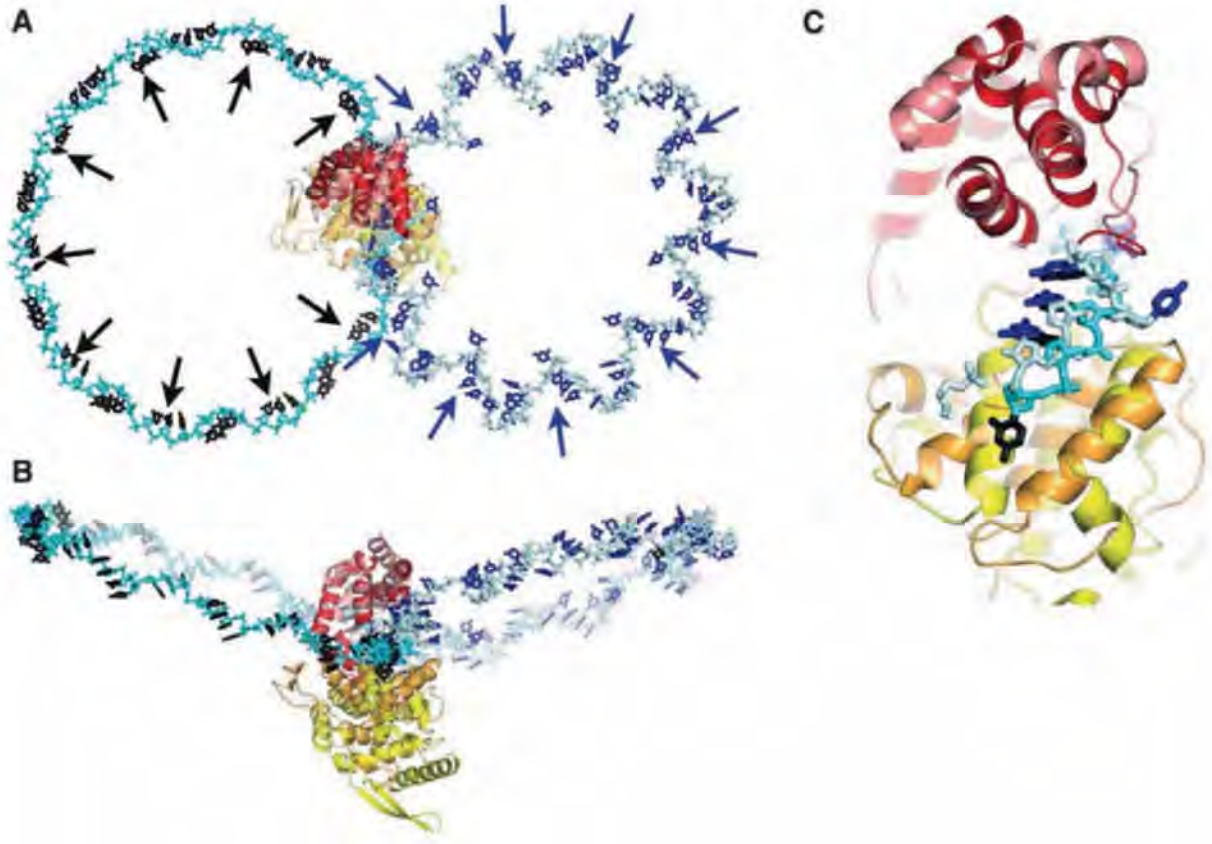


Fig. 3. Comparison of RSV with the VSV decameric ring. Overlap of the RNA rings (RSV on the left, 70 bases, and VSV on the right, 90 bases) after superposition of the conserved domains (red and yellow in Fig. 2) of one N subunit. The arrows point to the sites where each N protein binds three stacked bases in each ring. Blue and black arrows indicate the binding sites of VSV and RSV N, respectively. (**A**) Top view, (**B**) side view. (**C**) Close-up showing the remarkable superposition of the three bases buried within the N protein groove, considering that in VSV, an intervening base is looped out. In the last three panels, the bases in the VSV RNA are blue instead of black, for clarity; VSV N is colored pale red and orange for the CTD and NTD, respectively (N and C arms were removed for clarity).

of nucleocapsid-like helical assemblies of recombinant N complexed with cellular RNA (Fig. 4B). The reconstruction showed that the repeating units form lateral contacts resembling those observed in the ring. Because of the limited resolution, we did not attempt a direct fitting of the individual N subunits into the EM reconstruction, but we used the least possible distortion to the ring contacts to generate the corresponding helix, as described in (*13*). The packing of the subunits in the ring indeed suggests a simple way of modeling the helical nucleocapsid with minor slippage about the lateral contacts. The model showed that the NTD is easily recognized as forming the spikes projecting at roughly 45° from the nucleocapsid axis (Fig. 4, C and D). The loose contacts between N subunits in the ring can readily adapt to the distortion introduced by enforcing a helical axis instead of the 10-fold ring axis, while maintaining the RNA connectivity, as shown in movies S1 to S3. The RNA can easily follow the contacts, making a helix of a pitch varying between 69 Å (which is the minimum pitch possible to avoid clashes with the subsequent turn) to more than 100 Å, with 10 to 11 N proteins per turn, accounting for the observed flexibility of the nucleocapsid. Also, the region between helical turns is occupied by the mobile C arm, which may play a functional role by providing added flexibility to the nucleocapsid. RSV P was shown to bind the C arm of N (*20*), an interaction that may allow the polymerase complex to distort the helical conformation of the nucleocapsid during RNA synthesis.

Our model for the RSV nucleocapsid reveals a plausible way for the polymerase to thread through the RNA, reading the bases without needing to disassemble the nucleocapsid helix. Indeed, the domain organization of N suggests that the polymerase can induce a hinge movement of the NTD with respect to the CTD. The elongated NTD would act as a lever, with the polymerase contacting at its distal end (Fig. 2, orange) and causing the hinge movement, which would result in a transient opening of the groove during RNA readout. The location of the resistance mutations to the RSV604 compound (*17*) are thus quite likely to point to an interaction site of N with the polymerase complex. The hinge movement can make the three buried bases flip out, resulting in 11 bases in a row available for readout (5671-234-5671, in the numbering of Fig. 1). Support for this interpretation comes from studies using a phosphorylation mutant of RSV P that is impaired in transcription elongation, leading to the accumulation of abortive transcripts between 9 and 11 nt long (*21*). Given the large size of the polymerase complex, it is plausible that during elongation, the complex can easily maintain at least 3 or 4 consecutive N subunits [as suggested for rhabdoviruses (*22*)] in an open-hinge conformation, allowing the release of 21 to 28 nucleotides from the nucleoprotein grip, such that these nucleotides can be accommodated within the polymerase active site for synthesis of the complementary strand. In this model, the N protein would act as a helicase, dissociating the transient double-stranded RNA segment during procession of RNA synthesis along the genome.

Several studies of RSV RNA synthesis [reviewed in (*5*)] indicate that both the structure of the 3′ end of the RSV nucleocapsid, together with the specific 3′ terminal RNA sequence, are recognized by the viral polymerase for initiation of viral RNA synthesis. Our model shows that the 3′ promoter sequence and the first gene-start signal for transcription are spatially close to each other, at the first turn of the nucleocapsid helix (Fig. 4D).

The structure of the RSV RNP ring, together with the derived atomic model of the helical nucleocapsid, suggests important common features of the template for transcription and replication of viruses in the *Mononegavirales* order, which also includes other human pathogenic viruses such as those causing measles, mumps, Ebola

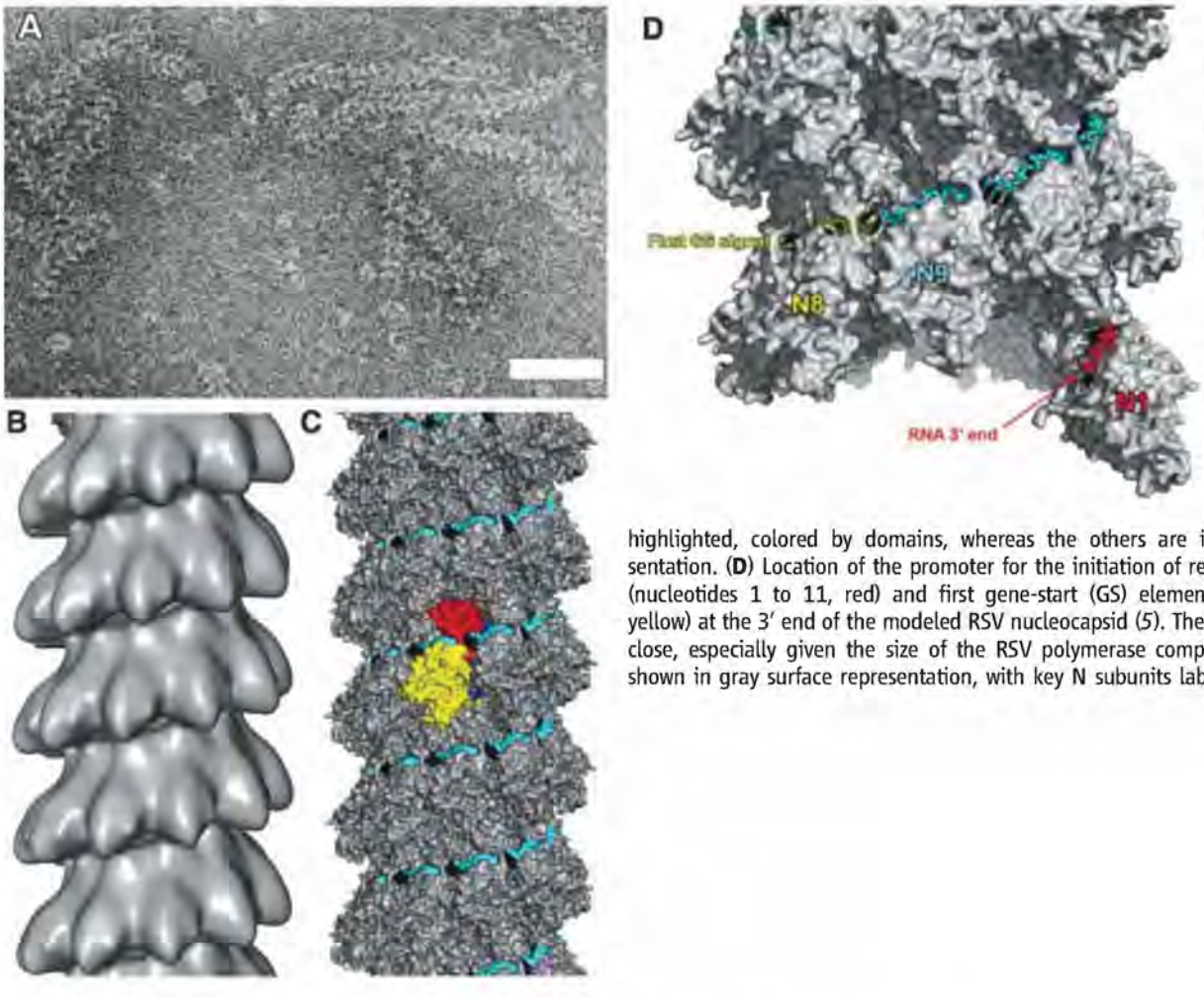


Fig. 4. The helical nucleocapsid. (**A**) Cryogenic negative-stain electron micrograph of recombinant RSV nucleocapsid-like helices. Scale bar, 50 nm. (**B**) 26 Å–resolution 3D reconstruction calculated from cryogenic negative-stain images. The helix comprises 9.8 N subunits per turn and has a pitch of 69 Å. (**C**) The helical nucleocapsid modeled from the contacts in the ring, using the same pitch (fig. S1), which results in 10.35 N subunits per turn (*13*). One subunit is highlighted, colored by domains, whereas the others are in gray, in surface representation. (**D**) Location of the promoter for the initiation of replication and transcription (nucleotides 1 to 11, red) and first gene-start (GS) elements (nucleotides 45 to 54, yellow) at the 3′ end of the modeled RSV nucleocapsid (*5*). The two sites are spatially very close, especially given the size of the RSV polymerase complex. The protein moiety is shown in gray surface representation, with key N subunits labeled from the 3′ end.

fever, and rabies. Furthermore, in the case of RSV, these results reveal important interaction sites, for instance, the cavity for three stacked bases, the site of insertion of the N arm, or the tip of the NTD where the resistance mutations arise, which can be specifically targeted for the development of therapeutic treatments, interfering with encapsidation or other roles of N.

References and Notes

1. P. L. Collins, J. E. Crowe, in *Fields Virology* (Lippincott, Philadelphia, PA, ed. 5, 2007), pp. 1601–1646.
2. M. P. Frogel, *Manag. Care* **17**, 7 (2008).
3. C. R. Pringle, *Order Mononegavirales*, in *Virus Taxonomy* (Elsevier/Academic Press, London, 2005), pp. 609–614.
4. A. A. Brunt *et al.*, *Family Paramyxoviridae*, in *Virus Taxonomy* (Elsevier/Academic Press, London, 2005), pp. 655–671.
5. V. M. Cowton, D. R. McGivern, R. Fearns, *J. Gen. Virol.* **87**, 1805 (2006).
6. S. M. Horikami, J. Curran, D. Kolakofsky, S. A. Moyer, *J. Virol.* **66**, 4901 (1992).
7. K. R. Qanungo, D. Shaji, M. Mathur, A. K. Banerjee, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5952 (2004).
8. T. Bachi, C. Howe, *J. Virol.* **12**, 1173 (1973).
9. D. Bhella, A. Ralph, L. B. Murphy, R. P. Yeo, *J. Gen. Virol.* **83**, 1831 (2002).
10. M. G. Rudolph *et al.*, *Structure* **11**, 1219 (2003).
11. T. J. Green, X. Zhang, G. W. Wertz, M. Luo, *Science* **313**, 357 (2006).
12. A. A. Albertini *et al.*, *Science* **313**, 360 (2006).
13. See supporting material on *Science* Online.
14. K. Maclellan, C. Loney, R. P. Yeo, D. Bhella, *J. Virol.* **81**, 9519 (2007).
15. K. El Omari *et al.*, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64**, 1019 (2008).
16. L. Holm, S. Kaariainen, P. Rosenstrom, A. Schenkel, *Bioinformatics* **24**, 2780 (2008).
17. J. Chapman *et al.*, *Antimicrob. Agents Chemother.* **51**, 3346 (2007).
18. M. Juozapaitis, M. Coiras, J. Staniulis, K. Sasnauskas, *Biologija (Vilnius)* **3**, 79 (2006).
19. A. A. Antson *et al.*, *Nature* **401**, 235 (1999).
20. J. Murray, C. Loney, L. B. Murphy, S. Graham, R. P. Yeo, *Virology* **289**, 252 (2001).
21. L. C. Dupuy, S. Dobson, V. Bitko, S. Barik, *J. Virol.* **73**, 8384 (1999).
22. A. Albertini, G. Schoehn, W. Weissenhorn, R. Ruigrok, *Cell. Mol. Life Sci.* **65**, 282 (2008).

23. This work was initiated as part of the sixth European Union Framework Programme consortium (http://cordis.europa.eu/life/src/control/qlk2-ct-2001-01225.htm). We thank A. Albertini, J. Bernard, D. Gerlier, A. Haouz, J. Lepault, M. Moudjou, C. Schulze-Briese, E. Stura, P. Weber, and R.P. Yeo for help and/or discussion; M. Backovic, D. Kolakovsky, J. Melero, L. Roux, and A. Tortorici for comments on the manuscript. Diffraction data collection was done at beamline X06SA of the Swiss Light Source and ID23-2 of the European Synchrotron Radiation Facility. R.G.T. benefits from a Marie-Curie RTN fellowship (consortium MRTN-CT-2006-035599, acronym EIHCV). This work was also supported by the French Fondation pour la Recherche Médicale (postdoctoral fellowship to P.F.V.), by a grant from the French Agence Nationale pour la Recherche (Programme MIME) to F.A.R., and by Merck-Serono. The coordinates and structure factors of the N-RNA rings were deposited in the Protein Data Bank with accession number 2wj8. The 3D EM reconstruction was deposited in the Electron Microscopy Data Bank with accession code EMD-1622.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1279/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S3
References
Movies S1 to S3

12 June 2009; accepted 8 October 2009
10.1126/science.1177634

UNLOCKING THE SECRETS OF microRNA

Microarrays are a vital tool in research on microRNAs, the tiny noncoding stretches of genetic material that regulate messenger RNA. They can't be beat for reading the microRNA "fingerprints" unique to disease states. But keeping up with this super hot field is a challenge for commercial microarray makers, and competition is fierce.
By Anne Harding

Carlo M. Croce may just have the world's most-tested and best-validated microRNA microarray. At six years and counting, it's certainly the oldest.

Croce, chairman of the department of molecular virology, immunology and medical genetics at **The Ohio State University** in Columbus, was the first investigator to show aberrant microRNA expression in human cancer, in 2002. The following year, he and his colleagues developed the first chip to detect global expression of microRNA. They've been using and upgrading it ever since.

The latest iteration of Croce's chip carries about 700 validated human and mouse microRNA and their precursors, as well as all the known ultraconserved noncoding RNAs. To date, Croce and his team have used the chip to test tissue from some 12,000 cancer patients. "I think that our chip might be the most tested chip of all," says Croce.

Microarray manufacturers have been racing to keep pace with this white-hot and speedily evolving field, which exploded after Croce's discovery that two microRNA genes targeting BCL2 showed altered expression in most cases of chronic lymphocytic leukemia (CLL). Most crucially, they've had to develop innovative ways to keep their products up-to-the-minute in a field where the latest thing today can be old news tomorrow. **Agilent Technologies** started making microarrays for studying microRNA in May 2007, when there were about 470 microRNAs in miRBase, a searchable database of all published miRNA sequences. As of September 2009, there were over 10,000 validated sequences in the database.

Powerful micro Managers

MicroRNAs are small stretches of genetic material—around 20 nucleotides long—that regulate gene activity by acting on messenger RNA. At the moment, cancer detection and diagnosis seem to be their most promising clinical application, although other burgeoning research areas include neurology and cardiovascular disease. And microarrays are a key tool for identifying the microRNA expression "fingerprints" or "signatures" of various disease states.

Nearly every human tumor type has been shown to have some alteration of microRNA expression, notes Dan Gallahan, deputy director of the Division of Cancer Biology at the **National Cancer Institute**. "What we're starting to see is that specific cancers have specific signatures," he adds. "Currently, the best way for us to generate those signatures is through microarrays."

MicroRNA researchers typically work with a mix of home-made and commercial arrays. Croce also uses commercial microRNA microarrays, to do parallel experiments and validate results. Some companies will do it all for you, like **LC Sciences**, which was among the first firms to develop microRNA analysis services, back in 2005. (**Invitrogen** and **Exiqon** entered the market that year, too.)

CREDIT: © ISTOCKPHOTO.COM/CGLADE

> "One microRNA may affect hundreds of genes. It's really the unraveling of this network that is the challenge ahead."

Look for these **Upcoming** Articles

High Content Screening — February 19

Genomics 1: Preparative Technologies for Next-Gen Sequencing — April 9

Fluorescent Labels — April 16

Inclusion of companies in this article does not indicate endorsement by either AAAS or Science, *nor is it meant to imply that their products or services are superior to those of other companies.*

Such omnibus offerings are ideal for researchers looking to test the waters of microRNA research without a major investment of time or resources. "The service model has worked really well for this field because it is so new and changing so rapidly," says Chris Hebel, vice president of business development at LC Sciences, which is based in Houston, Texas.

While there are drawbacks to using microarrays for studying microRNAs, including specificity issues and the risk of cross-hybridization because many mature microRNA sequences are so similar, researchers in the field agree that they will remain a key tool for conducting microRNA research for the foreseeable future. And companies that make microRNA arrays are doing their best to work out the kinks.

"I certainly believe there's still a very important place for these kinds of profiling studies," says Joshua Mendell, an associate professor of genetics at **Johns Hopkins University School of Medicine** in Baltimore, Maryland, and an early career scientist of the Howard Hughes Medical Institute. "To have a global view of microRNA dysregulation," says Croce, "the chip is still the way to go."

Giving Customers What They Want—Fast

"This is such a rapidly evolving market with a lot of new things coming to light, a lot of interest," says Sangita Parikh, senior project manager for microRNA and gene expression at Agilent, in Santa Clara, California. "Most people don't really know what's going to happen with microRNA, how useful it's going to be. Until the picture clears up, it will be quite difficult for the platform providers to provide a solution that the customer needs."

For now Agilent is focusing on keeping up by making its microarray tools and related products customizable and easy to use. The company bases its content on miRBase, and has a tool called the E-Array that allows users to tap into this database and design their own arrays. "The biggest thing is the customization that we offer, the flexibility to create your own design as you see fit," says Parikh.

After September's release of miRBase 14 (a jump to 10,883 sequences from May 2009's version 13, with 9,539), LC Sciences and Heidelberg, Germany–based **febit** jostled to be the first to announce their updated arrays; both put out press releases touting their new products on September 15.

But the jury's still out on just how many human microRNAs there are, and whether all the small RNAs being identified by deep sequencing (which involves sequencing every single small RNA in a sample) are worth studying. Mendell points out that the hairpin structure that is microRNA's calling card occurs fairly often, in the large amount of RNA that is produced by the genome. Such structures can theoretically be processed by the microRNA machinery and produce low abundance, nonfunctional small RNAs that resemble microRNAs. "I don't think we know how to distinguish between that class of molecules and the class of molecules that are actually functional, through sequencing alone," he adds.

Sorting this out, Mendell says, will be a key focus of microRNA research in the future.

Electronics, Microfluidics, and Tailored DNA

MicroRNA microarrays vary in the technology they use for probe capture, the substrate lying under these probes, the number of samples they can carry, and much more. Clearly, the ability to quickly incorporate newly released sequences into their microarray is an important bragging point for some manufacturers. Customization of arrays is another, as is the ability to work with a wide range of sample types—for example formalin-fixed, paraffin embedded (FFPE) material—and a variety of sample sizes.

"To have a global view of microRNA dysregulation, the chip is still the way to go"

The size of microRNAs means it's much easier to fish out intact sequences from tiny, and possibly degraded, samples—like FFPE—than trying to hunt down longer strings of genetic material. This is especially important since cancer diagnosis, prognosis, and treatment are such important potential clinical applications of microRNA profiling, and banked tumor biopsy samples are already proving to be an essential source of information.

"There are banks of these tumor tissues that are just begging for analysis," says Amy Cuneo, project manager for Invitrogen molecular biology reagents at **Life Technologies** in Carlsbad, California. "We have pretty robust protocols for using these sorts of tissue." Other companies also make microarrays suitable for testing FFPE samples for microRNA expression, including LC Sciences, Exiqon, and Agilent.

Both LC Sciences and febit use microfluidic chips, which make their rapid updates possible. Invitrogen's microarrays are printed on epoxide-coated glass. **CombiMatrix** bases its arrays on a semiconductor chip, and each feature on the chip is a digitally controlled microelectrode. According to CombiMatrix CEO and President Amit Kumar, the technology allows customers to easily update chips as their research progresses. "They can build a chip with whatever content they choose, they can try it out, and based on the results of the experiment, they can iterate it," Kumar explains. "Customers do that with our platform over and over again until they've got the chip they want to use in their study or for a diagnostic application."

The ability to read arrays electrochemically—with no need for optics, alignment requirements, or fluorescence requirements—means CombiMatrix can make very inexpensive microarray readers to use with its products. While a typical array reader goes for some $80,000 to $250,000, Kumar says his company sells one for $20,000 that's about the size of a toaster. CombiMatrix, which is headquartered in Mukilteo, Washington, is planning to release a reader that's the size of two cell phones put together, priced at $2,000.

Denmark's Exiqon has built a leading market position on its Locked Nucleic Acid (LNA) technology. The bond between DNA probes and microRNAs can be unstable, because the DNA monomer may move back and forth between the N conformation—ideal for binding to complementary RNA—and the S conformation. LNAs are DNA analogs in which a methylene bridge has been used to continued >

CREDIT: S. MAERK

lock the sugar ring, holding the monomer in the N position. LNA/RNA duplexes are stable at higher temperatures than DNA/RNA duplexes. And, according to the company, while the difference in melting point between sequences with no mismatches and one mismatch is 10 degrees centigrade when using DNA probes, it's 25 degrees with LNA probes.

This means Exiqon's microarrays are currently the most sensitive on the market, with 90 percent of probes capable of detecting concentrations below 10 attomolar, and allowing for microRNA profiling with as little as 30 nanograms of total RNA.

Exiqon sells a chip with the "428 unique human microRNAs" on its miRPlus chip, and another with 4,168 "mature microRNAs" from 85 species. Researchers can also buy probes from the company and put them on chips themselves.

More Samples, or More Probes?

Several companies have licensed Exiqon's LNA technology. **Luminex Corporation** used LNA in its first FlexmiR microRNA products, but has recently changed to a new approach that no longer relies on LNA. The new Luminex technology is capable of analyzing 100 different miRNA in a single microtiter well, with no need for sample labeling or amplification. Biotinilated capture probes are placed on microscopic polystyrene beads. After capture of the miRNA, an enzyme gets rid of unbound biotin, and will cleave off any mismatches, too. "Not only do we not see signal when there's no target, we don't see signal when there's a single-base mismatch," says Keld Sorensen, director of R&D at the Austin, Texas–based company.

The multiplexing capabilities of Luminex technology are useful for researchers focusing on up to 100 specific microRNAs who want to run many samples at once, Sorensen says. Luminex makes these liquid arrays to order.

"Every technology has a sweet spot," Sorensen says. "When you look at a printed array, it's the ability to get 1,000 microRNAs on a slide." For Luminex's technology, he adds, the sweet spot is the ability to run many, many samples at once, with each sample being tested for up to 100 miRNAs.

Also, Sorensen says, Luminex's products are particularly good for those studying plants. "People who are working with plants always feel like they are getting the leftovers," he says. "Our assay actually will work just fine with plant microRNAs as well."

Life Technologies' sweet spot may be its focus on offering a complete work flow for microRNA analysis, from a single-step process for isolating total RNA from cells or tissue to data analysis with its NCode Profiler Software (free of charge to customers), and miRNA qRT-PCR kits for data validation.

Cost effectiveness is, of course, a selling point too. **Phalanx Biotech Group** includes a cost estimating calculator with its miRNA Gene Expression Profiling Service, and boasts its ability to offer analysis for as little as $350 per sample. It also offers special deals on its website. Exiqon sells a startup package including three slides and a labeling kit for under $1,000, while the Invitrogen Ncode Human miRNA Array and **Miltenyi Biotec's** miRXplore Microarray Kits feature a sub-$2,000 startup cost.

SABiosciences has developed a cost-effective strategy for microRNA analysis that is microarray-like, but relies on polymerase chain reaction and does something current microarrays can't: measure expression levels. The company builds PCR arrays including the 88 "most abundantly expressed and best characterized microRNA sequences," in 96- and 384-well formats. It also sells pathway-focused arrays in human or mouse form, for cancer, cell differentiation and development, apoptosis, and more.

And like Invitrogen, SABiosciences provides a complete workflow, from tools for isolating samples to microRNA tailing kits.

Sorting Out Downstream Effects

As the field of microRNA research matures, understanding downstream effects is becoming increasingly important. Companies are also working on helping investigators take the next step to find microRNA targets. LC Sciences can tile the gene sequence of interest at single-base resolution across a chip. "It's such a complex network," notes Hebel. "One microRNA may affect hundreds of genes. It's really the unraveling of this network that is the challenge ahead."

Free target prediction programs are available to help researchers sort out the complexity, including **Memorial Sloan-Kettering Cancer Center's** miRanda, TargetScan, and PicTar. And many companies that make products for microRNA analysis have moved into the business of prediction as well; Agilent's GeneSpringGS helps researchers identify microRNA targets, look at connections with messenger RNA, and analyze pathways.

Hebel says LC Sciences is working on building microarrays that can probe for microRNAs and messenger RNA at the same time. "Having both of those profiles is going to go a long way toward having a diagnostic tool that might say you've got a cancer at a certain stage or not," he says.

Featured Participants

Agilent Technologies
www.agilent.com

CombiMatrix
www.combimatrix.com

Exiqon
www.exiqon.com

febit
www.febit.com

Invitrogen
www.invitrogen.com

Johns Hopkins University School of Medicine
www.hopkinsmedicine.org

LC Sciences
www.lcsciences.com

Life Technologies
www.lifetechnologies.com

Luminex Corporation
www.luminexcorp.com

Memorial Sloan-Kettering Cancer Center
www.mskcc.org

Miltenyi Biotec
www.miltenyibiotec.com

National Cancer Institute
www.cancer.gov

Ohio State University
www.osu.edu

Phalanx Biotech Group
www.phalanxbiotech.com

SABiosciences
www.sabiosciences.com

Anne Harding is a freelance science and medical writer based in Maplewood, New Jersey.

DOI: 10.1126/science.opms.p0900049

Muscular Dystrophy Array

The Cytosure Duchenne Muscular Dystrophy (DMD) array features a 4 x 44k format and dense probe coverage of the DMD gene region. It offers confidence in detecting deletions and duplications within the DMD gene. The CytoSure DMD array has undergone a process of empirical testing and optimization to provide probe sets of high sensitivity and specificity. Its average exon probe spacing of 10 bp (106 bp within introns) ensures excellent resolution. As a result, the entire DMD gene is covered on a single 44,000 feature array, which enables four full arrays per slide, maximizing cost efficiency by reducing the cost per sample.

Oxford Gene Technology
For information +44-(0)-1865-856826
www.ogt.co.uk

Wellplate for Nanoscale Volumes

The Mitos Wellplate holds small reagent droplets or small groups of cells. With a 40 x 40 array of nanoliter wells, each standard plate has a well volume of 0.26 nl to enable experiments to occur using nanoscale volumes. Covering a small area of 10 x 10 mm, the array of wells is designed to fit onto a microscope slide and under a standard coverslip. The wet etching process provides a surface of roughness of less than 5 nm, enabling easy access for microscopy systems. With a depth of 2 mm, the wellplate provides excellent optical transmission for clear observation of each well. Because each well can be etched to a range of depths and diameters, the well can be customized to offer alternate volumes. For more complex geometries, two layers can be used.

Dolomite
For information +44-1763-242491
www.dolomite-microfluidics.com

RNA Isolation

The new RNAzol RT is a single-step method that isolates total RNA with messenger RNA and small RNA (10–200 bases) in separate fractions. It provides higher RNA yield and quality than previous single-step reagents. There is no chloroform-induced phase separation. The user just adds water to remove DNA, proteins, polysaccharides, and other contaminants. The isolated RNA is ready to use in reverse transcription polymerase chain reaction and other applications, with no deoxyribonuclease treatment necessary. All steps are performed at room temperature.

Molecular Research Center
For information 800-462-9868
www.mrcgene.com

Blood Dry Storage Products

QIAsafe DNA Blood Tubes and Plates offer high DNA yield, easy transport, and storage at room temperature. The QIAsafe DNA Blood Kit provides easy transport and storage of whole blood at room temperature. The QIAsafe products offer the first dry blood storage method available on a matrix; the QIAsafe matrix is a mixture of dissolvable compounds that stabilizes DNA at room temperature. The technology is based on anhydrobiosis, a biological mechanism employed by some multicellular organisms to enable their survival in a dry state for periods of more than 100 years. In QIAsafe blood products, this technology is used to preserve DNA for future extractions from blood.

Qiagen
For information 240-686-7660
www.qiagen.com

Human Transcription Factors

Chromatin immunoprecipitation (ChIP) to identify transcription-factor binding sites depends on monospecific antibodies to the transcription factors of interest, but for many transcription factors such antibodies are not available. To overcome this limitation, SpectraGenetics now offers more than 100 genomic clones that express human transcription factors with 3xFLAG tags at their C-termini. Transcription factors are expressed at physiologically normal levels in transfected cells and the normal set of alternatively spliced isoforms is represented. ChIP analysis follows standardized protocols that use anti-FLAG antibodies for immunoprecipitation.

SpectraGenetics
For information 412-488-9350
www.spectragenetics.com

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availabilty of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

POSITIONS OPEN

Yale

The Section of Cardiovascular Medicine seeks exceptional candidates to establish research programs in cardiovascular developmental biology and genetics. Successful individuals will have Ph.D. and/or M.D. degrees and will have a proven record of originality and productivity. The Section is undergoing a major expansion and houses a newly formed Yale Cardiovascular Research Center. Yale University School of Medicine has established a close affiliation with University College London, United Kingdom, that among other benefits provides access to extensive clinical and translational genetics resources at UCL. Although all extraordinary candidates will be considered, we are especially interested in **DEVELOPMENTAL BIOLOGISTS** utilizing zebrafish as their model, who will be able to establish and operate a zebrafish research facility, and in **GENETICISTS** involved in translational genetics research, who will be able to run or participate in a clinical cardiovascular genetics program in addition to basic research efforts.

Please electronically send your curriculum vitae with a list of publications, a summary of research (two pages), and a research plan (three pages) along with the names of three references by December 1, 2009, to **e-mail: michael.simons@yale.edu.**

ASSISTANT/ASSOCIATE PROFESSOR
Animal Physiology

A tenure-track opening for an **ANIMAL PHYSIOLOGIST** is available in the Department of Biological Sciences at DePaul University starting September 2010 (pending approval of funding). The successful candidate will be broadly trained in animal physiology with a strong commitment to undergraduate education. We welcome endocrinologists, neurobiologists, and organismal-based physiologists, but all subdisciplines and animal model systems will be considered. Ph.D. required; postdoctoral and previous teaching experience preferred. Teaching responsibilities to include some combination of: introductory biology for nonmajors; co-teaching one-quarter of introductory biology for majors; vertebrate physiology; and graduate/advanced undergraduate courses in the candidate's area of expertise. The Department is housed in a spacious and well-equipped teaching, research, and support complex that includes a 2,000 square foot staffed animal care facility and a newly opened gold-LEED certified wing used by biology, chemistry, and environmental science. Startup funds are provided. Review of applications will begin December 15, 2009, and will continue until the position is filled.

Applicants should apply online with a cover letter; curriculum vitae; a two-page statement of research and teaching interests, including teaching philosophy; and a general list of equipment and supply needs. In addition, three letters of reference should be sent to: **Animal Physiology Search Committee, Department of Biological Sciences, DePaul University, 2325 N. Clifton Avenue, Chicago, IL 60614.** Additional inquiries should be sent to the Search Committee Chair, **Dr. Elizabeth LeClair, e-mail: eleclair@depaul.edu**.

To apply, please visit the following **website: facultyopportunities.depaul.edu/applicants/Central?quickFind=50642**.

The Department seeks diversity in its faculty and we encourage applications from women, people of color, and the members of other historically underrepresented groups. DePaul University is committed to diversity and equality in education and employment.

An NIH-funded **POSTDOCTORAL POSITION** is available in the field of obesity and type 2 diabetes. Background in the field is required. Experience with clamp study, adenovirus construction, adipose tissue biology, and inflammation is also required. Salary is consistent with NIH guidelines. Interested candidates with Ph.D or M.D. degree please submit curriculum vitae to: **Haiyan Xu, M.D.-Ph.D., Brown Medical School, 55 Claverick Street, Room 318, Providence, RI 02903. E-mail: haiyan_xu@brown.edu.**

POSITIONS OPEN

MOLECULAR DEVELOPMENTAL BIOLOGIST

Tenure-track **ASSISTANT PROFESSOR** position beginning fall 2010 in the Indiana University Northwest Department of Biology. Primary responsibilities include teaching upper-level undergraduate developmental, molecular, and related biology courses. Ideal candidates will be expected to demonstrate a commitment to teaching that includes undergraduate research.

Qualifications: Ph.D. is required; postdoctoral research experience is preferred. Expertise in molecular and/or evolutionary development and interest in biotechnology and bioinformatics are desirable.

Submit curriculum vitae, unofficial graduate and undergraduate transcripts, three letters of recommendation, and a statement of teaching and research philosophy that explains integration of teaching and research to: **Biology Search Committee, Department of Biology, Indiana University Northwest, 3400 Broadway, Gary, IN 46408.**

Application review begins on January 4, 2010, and will continue until the position is filled. For more information on the Department, see the biology **website: http://www.iun.edu/~biologyn**.

One of eight campuses of Indiana University, IU Northwest is located in metropolitan northwest Indiana, approximately 30 miles southeast of Chicago and 10 miles from Indiana Dunes National Lakeshore. The campus has a diverse student population approaching 6,000 and offers Associate, Baccalaureate, and Master's degrees in a variety of undergraduate and graduate programs in arts and sciences, business and economics, education, nursing and health professions, public and environmental affairs, and social work.

Indiana University is an Equal Opportunity/Affirmative Action Employer committed to achieving excellence through diversity. The University actively encourages applications from women, minorities, persons with disabilities, and members of other underrepresented groups.

TENURE-TRACK FACULTY POSITIONS IN INFLAMMATION

The Department of Pathology, Microbiology and Immunology, School of Medicine, University of South Carolina, Columbia, invites applications for tenure-track **ASSISTANT PROFESSOR** positions in the broad area of inflammation. Candidates must have a Ph.D. or M.D. or equivalent with postdoctoral research experience. Outstanding applicants working in any area of inflammation complementary to our existing faculty (**website: http://pmi.med.sc.edu/default.html**) will be considered. The areas include inflammation as it relates to bacterial and viral infections, cancer, toxicology, and complementary and alternative medicine to treat autoimmune/inflammatory diseases. Candidates are expected to develop a strong, extramurally funded research program using cutting-edge cellular and molecular techniques as well as to participate in the teaching mission of the Department. Investigators who have shown evidence of independence through active grant funding are encouraged to apply. Competitive salary and startup funds are available. Apply with curriculum vitae, statement of research plans, and three letters of recommendation to: **Dr. Mitzi Nagarkatti, Chair, Department of Pathology, Microbiology and Immunology, University of South Carolina School of Medicine, Columbia, SC 29208. Or e-mail: inflammation@uscmed.sc.edu.** The search will start immediately and continue till the positions are filled. *USC Columbia is an Equal Opportunity/Affirmative Action Employer and encourages applications from women and minorities.*

RESEARCH ASSOCIATE in Theoretical Atomic, Molecular, and Optical Physics

Applications are invited for one or more **POSTDOCTORAL** openings in the group of **Professor Brett Esry** at Kansas State University to study ultracold few-body collisions. Applicants must have a Ph.D. (or equivalent) in physics or chemistry and should have a strong background in theoretical atomic or molecular physics. For more information and application procedures, see **website: http://www.phys.ksu.edu/personal/esry**. *Equal Employment Opportunity Employer. Background check required.*

WWW.NIH.GOV
Positions @ NIH
THE NATIONAL INSTITUTES OF HEALTH

Staff Scientist, Laboratory of Immunology

The National Institute of Allergy and Infectious Diseases (NIAID) conducts and supports a global program of research to better understand, treat, and ultimately prevent infectious, immunologic, and allergic diseases. NIAID's Laboratory of Immunology, Lymphocyte Biology Section (LBS)/Program in Systems Immunology and Infectious Disease Modeling (PSIIM) has an opening for a staff scientist with expertise in optical imaging.

LBS/PSIIM's aim is to create a detailed understanding of immune responses to foreign pathogens or self antigens, as well as to develop new tools for predicting how the immune system will respond if perturbed, for example, by a vaccine. The laboratory uses optical imaging involving both 2-photon and confocal instruments in pursuing these goals (www.niaid.nih.gov/labs/aboutlabs/li/lymphocyteBiologySection/germain.htm).

The successful applicant will have primary responsibility for maintaining and operating four distinct 2-photon instruments (a custom dual laser, resonant scanner device; a Zeiss 510; a Zeiss 710; and a BioRad radiance 2100), two of which are also fully equipped confocal instruments. The successful applicant will work with LBS/PSIIM postdoctoral fellows in conducting their imaging studies as well as with staff from collaborating laboratories. Resources will be available to upgrade and modify these instruments, to acquire new instruments, and to develop novel imaging probes and methods. Software development will be encouraged, especially in collaboration with PSIIM computational staff.

Additionally, there will be opportunity for the successful applicant to design, conduct, and publish the results of biological imaging experiments. Strong technical applicants lacking an immunology background who are otherwise excellent fits to the imaging needs will be given full consideration. If hired, such an applicant will be provided opportunities for training and education in immunology and infectious diseases to facilitate development of a strong experimental research program in addition to technical activities and collaborative work with other members of LBS/PSIIM.

National Institute of Allergy and Infectious Diseases

To apply, send curriculum vitae, bibliography, three letters of recommendation, and a short statement about your skill set including how you see applying optical imaging to better understanding immunity (no more than two pages) by **January 31, 2010**, to Ronald Germain, M.D., Ph.D., Chief, LBS; Director, PSIIM, NIAID/NIH/HHS, 10 Center Drive, MSC 1892, Building 10, Room 11N311, Bethesda, MD 20892-1892 or rgermain@nih.gov.

For more information about NIAID and to view additional career opportunities, please visit **www.niaid.nih.gov/careers/ssds.**

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

National Institute of Allergy and Infectious Diseases

Proud to be Equal Opportunity Employers

THE NIH IS DEDICATED TO BUILDING A DIVERSE COMMUNITY IN ITS TRAINING AND EMPLOYMENT PROGRAMS

UNIVERSITY OF LIVERPOOL

Department of Chemistry

Postdoctoral Researchers

£30,594 - £32,458 pa (under review)

Nanoporous Materials (4 Posts)

Four 3-year positions in the synthesis and sorption/reactivity characterisation of nanoporous inorganic, hybrid and organic materials are available. Expertise in porous materials, excellent publication record and a PhD in chemistry or materials science are essential. Experience in organic synthesis (eg, of ligands or polymers), sorption measurements, catalysis or X-ray diffraction is desirable. The posts will form part of a diverse 5-year programme in materials discovery, with one of the posts being the theme leader. You should combine specific high-level technical skills in porous solids with an interest in the broader, cross-disciplinary application of such materials. **Job Ref: R-570618/S**

Oxide Materials (3 Posts)

Positions in (1) oxide materials discovery; (2) the characterisation of the electronic properties of oxides; and (3) evaluation of oxide electroceramics are available for 3 years. Expertise in metal oxides (1-3), diffraction (1), magnetic and electronic characterisation of solids (2), and either ferro/piezoelectrics (3) or SOFC materials (3) is essential, together with a PhD in chemistry, physics or materials science. The posts will form part of a diverse 5-year programme in materials discovery. You should combine specific high-level technical skills in oxides with an interest in the broader, cross-disciplinary application of such materials. **Job Ref: R-570617/S**

Organic Synthesis For Materials Discovery (1 Post)

We seek an organic chemist to underpin the synthetic organic component of a major 5-year materials discovery programme. A track record in targeted synthesis, a PhD in chemistry, and an excellent associated publication record are essential. Experience of synthesis related to materials problems such as ligands and polymers would be an advantage. You will work within a large team of approximately 20 researchers, mostly with a materials background. This 3-year post will involve designing syntheses for target structures, the synthesis of these structures, and also advising and assisting members of the team with less experience of organic synthesis. **Job Ref: R-570622/S**

Nanomaterials (2 Posts)

Two 3-year positions in nanomaterials synthesis are available. Expertise in either polymer nanostructures or inorganic nanomaterials is essential, as evidenced by an excellent publication record and a PhD in either chemistry or materials. Experience in polymerisation, emulsions, nanoparticles, self-assembly, microscopy or colloid science is desirable. The posts will form part of a diverse 5-year programme in materials discovery, with one of the posts being the theme leader. You will combine specific high-level technical skills in nanomaterials with an interest in the broader, cross-disciplinary application of such materials in themes such as porous materials and functional oxides. **Job Ref: R-570608/S**

High Throughput Materials Discovery (1 Post)

We seek a materials chemist to underpin the high-throughput and combinatorial aspects of a major 5-year materials discovery programme. An excellent publication record in the synthesis of materials is essential. Direct experience of using high-throughput technology (eg, robotics, automation) or design of experiments (DoE) would be an advantage. Knowledge of organic synthesis, inorganic materials, porous materials, or nanomaterials would be an advantage. This 3-year post will be based in the Centre for Materials Discovery (www.materialsdiscovery.com) and will enable the development of accelerated, automated routes to the synthesis and characterisation of porous materials (eg, MOFs, organic polymers), oxides, and nanomaterials. **Job Ref: R-570615/S**

Closing date for all posts: 5 January 2010

For full details, or to request an application pack, visit www.liv.ac.uk/working/job_vacancies/

or e-mail jobs@liv.ac.uk Tel 0151 794 2210 (24 hr answerphone),
please quote Job Ref in all enquiries

COMMITTED TO DIVERSITY AND EQUALITY OF OPPORTUNITY

Penn Medicine

Faculty Position
Computational Biologist

The Penn Epigenetics Program and the Institute for Diabetes, Obesity, and Metabolism (IDOM, website: http://www.med.upenn.edu/idom/) at the University of Pennsylvania seek a highly qualified computational biologist for a faculty position as Assistant Professor in the tenure track or non-tenure research track. Academic track will be commensurate with credentials and experience. Applicants must have a Ph.D. degree or equivalent. The faculty appointment will be in the appropriate department in the School of Medicine.

We are seeking a computational biologist with research interests in the functional analysis of genome-wide datasets. An individual with background in the application or development of state-of-the art computational tools for evaluating genome-wide chromatin immunoprecipitation studies of transcription factors and epigenetic modifications/factors are most highly desired. The most appropriate candidate will take advantage of close interactions with colleagues in the Penn Epigenetics Program and the IDOM, which will be co-localized in a new research building in the Summer of 2010. Training in experimental biology and interest in systems biology, epigenetics/chromatin biology and/or metabolism would be highly useful, but not required.

Responsibilities will include participation in academic activities of Penn Epigenetics Program and the IDOM. To be considered for the tenure track, the successful candidate will be expected to develop an independently funded research program and will need to demonstrate excellent qualifications in teaching. The expected start date is July 2010, or after.

The University of Pennsylvania is an affirmative action/equal opportunity employer. Women and minorities are strongly encouraged to apply.

For Tenure Track, please apply to: http://www.med.upenn.edu/apps/faculty_ad/index.php/g323/d2137

For Research Track, please apply to: http://www.med.upenn.edu/apps/faculty_ad/index.php/g323/d2127

Research Complex at Harwell

MRC Medical Research Council

The Research Complex at Harwell (RCaH), Oxfordshire

Post-doctoral Research Scientists (x2) and Research Technician

Structural basis of Holliday Junction Resolution

Two postdoctoral posts and one technician post are available to work on a 5-year MRC-funded programme to determine structures of human protein complexes involved in the resolution of four-way DNA Holliday junctions. The Holliday junction is an important intermediate in DNA recombination and repair, and its cleavage is a critical step, since two of the four DNA strands must be cleaved correctly, at appropriate points, to free two daughter duplexes. The overall aim of the programme is to determine three-dimensional structures for resolution complexes and relate these to functional studies in collaborating laboratories to provide a complete mechanistic understanding of this fundamental biological process. Protein crystallography will be the primary method of structure determination, but will be coupled with complementary structural techniques, such as Electron Microscopy (EM), Small Angle X-ray Scattering (SAXS), Small Angle Neutron Scattering (SANS), Atomic Force Microscopy (AFM) to study higher order complexes in solution. A PhD in a relevant subject area and a strong research background in protein crystallography and protein chemistry are required for the postdoctoral positions, together with the proven ability to work independently.

The Research Complex at Harwell (RCaH) is an RCUK project, a joint initiative of the Research Councils and Diamond Light Source to provide new state-of-the-art multidisciplinary research laboratories. It is managed by the Medical Research Council (MRC) on behalf of the other partners, and will be completed by the end of 2009. RCaH will be equipped to the highest level for structural biology research, and benefits from its location adjacent to the Diamond Light Source, and close to the ISIS neutron source. RCaH also houses the Oxford Protein Production Facility-UK, the core CCP4 team and the STFC Lasers for Science Facility. More information is available at www.rc-harwell.ac.uk

For an informal discussion about the role please contact Dr. Steve Carr at stephen.carr@rc-harwell.ac.uk

Post Doctoral Research Scientist - Ref: RCH09/589

Research Laboratory Technician - Ref: RCH09/590

Benefits include 30 days annual leave, competitive salary, and a final salary pension scheme. Please note that final appointment will be subject to pre-employment screening.

Applications for these roles should be made on line at http://jobs.mrc.ac.uk If you do not have internet access or experience technical difficulties, please call 01793 301154 quoting the relevant reference number.

If you would like to receive this advert in large print, Braille, audio or electronic format/hard copy, please contact the Recruitment team at the MRC Shared Service Centre on the telephone number above or email recruitment@ssc.mrc.ac.uk

Closing date: 1 January 2010.

Interviews are planned for: early February 2010.

The MRC is an Equal Opportunities Employer

THE PEDIATRIC DENGUE VACCINE INITIATIVE
INTERNATIONAL VACCINE INSTITUTE

POSITION ANNOUNCEMENT
DIRECTOR

The Pediatric Dengue Vaccine Initiative (PDVI), a program hosted by the International Vaccine Institute (IVI), Seoul, Korea, is seeking a Director to oversee its development and operations. The PDVI has as its objective, to facilitate and accelerate the development and introduction of safe and effective dengue vaccines in developing countries. The PDVI has several programmatic components: 1) supportive research to facilitate the clinical testing of dengue vaccines by developing and improving immunological and diagnostic assays; 2) evaluation research focusing on development of field sites in dengue-endemic countries where clinical testing of dengue vaccines can take place and on field evaluations of dengue diagnostics and assays; 3) vaccine product development partnerships to achieve products designed for and affordable to dengue-endemic countries; and 4) activities to assure access to dengue vaccines, including research to better define the global dengue disease burden and disease costs, to model the potential cost-effectiveness of dengue vaccines, and to estimate the market for dengue vaccines, as well as collaborative activities with national and international partners, including the World Health Organization, to communicate this evidence and to plan for vaccine introduction into developing countries. The host organization, the IVI, is an international non-profit organization focused primarily on accelerating the research, development and introduction of new and improved vaccines for use primarily in developing countries. The PDVI headquarters are located at the IVI in Seoul, Korea.

The incumbent will be a recognized leader in one of the fields encompassed by the PDVI program, including virology, epidemiology, immunology, and vaccine development. A broad knowledge of dengue would be an advantage, as would experience working in the vaccine industry. The incumbent should have experience in program-building, resource mobilization, and in staff development.

Minimum qualifications include a doctorate degree in a relevant discipline, and significant experience in leading a multi-disciplinary field/ laboratory research program.

Salary will be internationally competitive. The Institute provides appropriate fringe benefits including a housing allowance, home leave, and income tax reimbursement.

The International Vaccine Institute is an independent international organization established under the Vienna Convention of 1969. It is governed by a Board of Trustees the majority of whom are elected based on their personal capacity.

Applications should be sent to:

Ms. Eunsuk Kim
Human Resources Officer
International Vaccine Institute
San 4-8 Nakseongdae-dong
Gwanak-gu, Seoul, Korea
Tel: 82-2-872-2801 Fax: 82-2-872-2803
Email: eskim@ivi.int

From whom further particulars can be obtained. Absolute confidentiality will be respected.

NUS National University of Singapore

Positions in Environmental Biology

Department of Biological Sciences, National University of Singapore

The Department of Biological Sciences, National University of Singapore (NUS), invites applications for tenure-track faculty positions at the Assistant/Associate Professorship level in the following areas of Environmental Biology:

- Microbial ecology
- Ecotoxicology
- Microalgal systematics and ecology
- Environmental genomics
- Applied entomology
- Urban ecology
- Climate change biology
- Environmental modeling

An interest in tropical urban environments will be an advantage. Singapore also retains examples of several natural ecosystems and is at the centre of one of the biologically richest regions of the world. Candidates should have a Ph.D. with relevant postdoctoral experience, an outstanding publication record, and a strong commitment to both teaching and research. Appointment at the Professor level may be considered for exceptional candidates. Successful applicants will join a growing Environmental Biology group within a diverse and highly successful Department of Biological Sciences. The Department has 60 full-time faculty members and over 260 graduate students from 16 countries. Facilities and research support are world-class. NUS currently ranks number 20 in "Life Sciences & Biomedicine" in the Times Higher Education–QS World University Rankings. Please visit our website at www.dbs.nus.edu.sg/ for further details of the Department and its research and teaching programs.

Interested candidates should forward a letter of intent describing their career goals, research plans, teaching interests, curriculum vita and the names and addresses of four academic referees to:

Chair, Environmental Biology Search Committee (c/o Ms Lisa Lau)
Department of Biological Sciences
National University of Singapore
14 Science Drive 4
Singapore 117543
Fax: (65) 6779 5671; Email: dbsjobs@nus.edu.sg

Closing date of application: 28 February 2010

WAYNE STATE UNIVERSITY

Tenure/Tenure-Track Faculty Positions Available

The **Eugene Applebaum College of Pharmacy and Health Sciences** (EACPHS), Wayne State University (WSU), is seeking candidates for 12-month tenure/tenure-track positions at Assistant/ Associate/Full Professor levels in Pharmacy Practice, Pharmaceutical Sciences, Health Care Sciences-Physical Therapy and Clinical Laboratory Science.

Successful applicants will demonstrate excellence in research and teaching skills and will be expected to maintain a vigorous, externally funded research program. Qualifications are commensurate with rank but should include a doctorate in related field; licensure or eligible for licensure, if applicable; and a peer-reviewed publication record.

The Eugene Applebaum College of Pharmacy and Health Sciences is located on the Detroit Medical Center Campus, home to 11 hospitals, the WSU School of Medicine and the Detroit Veteran's Administration Hospital. Applicants can visit the college Web site at **www.cphs.wayne.edu** for more information on open faculty positions. Applicants also can view complete position descriptions and apply online at the WSU Human Resources Web page **www.hr.wayne.edu** under "Job Postings."

Wayne State University is among the nation's prestigious 3.6 percent of universities with Carnegie Classification of Research University, Very High research activity. WSU is home to more than 30,000 students and 350 degree programs. It is located in the heart of the University Cultural Center and a short distance from first-class museums, theaters, concert halls and sports arenas.

Wayne State University is an Affirmative Action/Equal Opportunity Employer.

Effective Lab Skills WEBINAR

MANAGING PEOPLE, PROJECTS, AND MONEY

December 2, 2009
12 noon Eastern Time
(9 a.m. Pacific, 5 p.m. GMT)

Academic scientists face many challenges when it comes to running their own lab. Beyond doing good research, lab heads need to consider many practicalities, including managing staff (hiring, retaining, motivating), developing reasonable timelines for projects and keeping them on schedule, and tracking grant budgets.

Watch this *Science* Careers live webinar to gain insight from experts related to effective laboratory management skills. Don't miss this opportunity to pick the brains of experienced and practiced academic managers.

Questions can be submitted live to the panel during the webinar or in advance via e-mail provided with registration. To register, visit **www.sciencecareers.org/webinar**

Participating Experts:

Karen M. Hill, Ph.D.
Assistant Professor
Howard University
Washington, DC

Klaus Nüsslein, Ph.D.
Associate Professor
University of Massachusetts at Amherst
Amherst, MA

Kelly Suter, Ph.D.
Associate Professor
University of Texas at San Antonio
San Antonio, TX
ScienceCareers.org Forum adviser

Produced by the
Science/AAAS Business Office.

Science Careers
From the journal *Science* AAAS

NYU School of Medicine

The Department of Surgery at NYU Langone Medical Center is seeking to recruit a Basic Scientist, for The Helen and Martin Kimmel Division of Wound Healing & Regenerative Medicine under the direction and leadership of **Harold Brem, MD**.

The Division of Wound Healing is seeking to fill several faculty positions at the rank of Assistant, Associate or Full Professor for scientists that use cellular and/or molecular-genetic approaches to address fundamental aspects of Wound Healing and Regenerative Medicine. The research laboratory is extremely collaborative with numerous resources in place. The program is purely translational and literal bridge to the operating room exists. Candidates with a background in wound healing, angiogenesis, pathology, stem cell biology, and regenerative cardiovascular medicine using mammalian systems or genetic approaches are especially encouraged to apply. Competitive salary, laboratory space and start-up funds are available. Candidates are expected to have a track record from the NIH. A PhD, MD, PhD/MD (or equivalent) is essential.

Website: www.nyuwound.org

Interested individuals should send their resume to the **Helen and Martin Kimmel Program Manager, Tracy Henry** via email to **tracy.henry@nyumc.org**.

NYU School of Medicine

Bioinformatics:
Applications are sought in Bioinformatics for NIH funded researcher to work with **Dr. Harold Brem**, Division Chief of the Helen and Martin Kimmel Wound Center at New York University School of Medicine Department of Surgery- Division of Wound Healing and Regenerative Medicine.

Responsibilities and Duties:
- Data Analysis
- Develop algorithms and software tools in support of project research objectives
- Participating in writing project updates, reports, proposals and scientific articles
- Correlate Wound Electronic Medical Record (WEMR) data to clinical outcomes
- Collaborating with investigators within and outside the program, in furtherance of project objectives

Requirements:
Applicants must hold a MD or PhD degree in Bioinformatics, Biology, Engineering, Physics, Computer Science, Mathematics or a related scientific discipline pertinent to the position. Experience with bioinformatics methods, tools, websites and data resources highly desired. Track record from NIH preferred.

Qualified candidates should send your CV to the **Helen and Martin Kimmel Program Manager Tracy Henry** via email to **tracy.henry@nyumc.org**.

OLD DOMINION UNIVERSITY

MICROBIOLOGY, IMMUNOLOGY OR MOLECULAR CELL BIOLOGY FACULTY POSITIONS

The Department of Biological Sciences at Old Dominion University is recruiting for mid to senior level faculty in the area of Microbiology, Immunology or Molecular Cell Biology. To be considered, investigators must have a strong record of research accomplishments as determined by: publications in top quality journals, peer reviewed grant support, and evidence of high quality teaching. Preference will be given to candidates whose research complements existing programs in the Department and/or College. The successful candidates are expected to have national/international recognition, strong record of publications and a research program consistently funded by peer-reviewed research grants. State support and a competitive start-up package will be provided. The successful candidates will be expected to contribute to our teaching of undergraduate students, as well as to develop graduate courses in their area of expertise.

Applications should include a curriculum vitae, a summary of research and teaching experience, statement of current research interests and future plans, as well the names, addresses, e-mail address, and phone numbers of 4 references. All applicants must have a Ph.D. and/or M.D. or equivalent. Please send applications and nominations to the **Search Committee (MICB) at MICB@odu.edu.** Review of applicants will begin immediately and continue until the positions are filled. The anticipated start date for these positions is July 2010. The Department receives substantial support from state funds and from research grants from federal and other agencies. The Department has strong graduate programs which currently have over 100 students, including a Ph.D. program in biomedical sciences and a Ph.D. program in ecology

Old Dominion University is an Affirmative Action/Equal Opportunity institution and requires compliance with the Immigration Reform and Control Act of 1986.

OSWEGO
STATE UNIVERSITY OF NEW YORK

ASSISTANT PROFESSOR

Plant Molecular Biologist

The Department of Biological Sciences at the State University of New York at Oswego invites applications for a full-time tenure-track position at the Assistant Professor level with expertise in plant molecular biology.

Required Qualifications: Ph.D. (or completion no later than the end of the Fall 2010 semester) with expertise in plant molecular biology; Commitment to excellence in teaching and providing meaningful learning experiences to students from diverse backgrounds. Applicants with experience in plant physiology, ecophysiology or plant microbial interactions will be preferred.

For more information about the department, see **www.oswego.edu/biology**. For more information about this position and application procedures, please see:

http://www.oswego.edu/vacancies

SUNY Oswego is an Affirmative Action Employer

The Department of Structural and Cellular Biology, Tulane School of Medicine, Tulane University Health Sciences Center, New Orleans, LA

Non-Tenure Track Faculty Position

The Department of Structural and Cellular Biology is seeking applications of all qualified candidates for a position (at the Associate Professor to Professor level) as the Assistant Director of Anatomical Teaching. This individual will assist in teaching and directing our Human Gross and Developmental Anatomy and Human Histology courses and will also assist in our Medical Neuroscience course. Candidates must have an earned doctorate (Ph.D., D.V.M., D.O., or M.D.). Prior experience in teaching of Neuroscience is preferred, but teaching experience in Human Gross Anatomy and Histology is essential. Documentation of excellent teaching skills is required. Extramural research funding is not required.

Applicants should submit a cover letter, curriculum vitae, summary of teaching experiences and philosophies and letters of references to: **Dr. Steven M. Hill, Department of Structural and Cellular Biology, SL49, Tulane University Health Sciences Center, School of Medicine, 1430 Tulane Avenue, New Orleans, LA 70112** by **December 15, 2009**.

Tulane University is an Affirmative Action/Equal Opportunity Employer. Women and minorities are encouraged to apply.

Cold Spring Harbor Asia Conferences

Suzhou Dushu Lake Conference Center, Suzhou, China

James Watson Cancer Symposium

April 6 - 11, 2010
Abstract deadline: January 15, 2010

Arranged by:
Scott Lowe (USA)
Tak Mak (Canada)
Yusuke Nakamura (Japan)
Karen Vousden (UK)
Xiaodong Wang (China / USA)

Keynote Speakers
Zhu Chen
Hans Clevers
Bruce Ponder
Tadatsugu Taniguchi
Craig Thompson

Topics:
- Signaling and Cancer
- Tumor suppressor networks
- Targeted therapeutics
- Tumor microenvironment and angiogenesis
- Tumor maintenance genes
- Molecularly stratified clinical trials
- Molecularly targeted cancer therapy
- Cancer genomics/epigenomics
- Stem cells and cancer
- MicroRNAs and cancer
- Immune surveillance

Invited Speakers
Hiroyuki Aburatani, James Allison, Alan Ashworth, Stephen Baylin, Lewis Cantley, Lynda Chin, Suzanne Cory, Sara Courtneidge, Ronald DePinho, Gerard Evan, Douglas Green, Nancy Jenkins, William Kaelin, Michael Karin, Narry Kim, David Lane, Jackie Lees, Yongzhang Luo, Hiroyuki Mano, Satoru Miyano, Kohei Miyazono, Sean Morrison, Luis Parada, George Peoples, David Sabatini, Charles Sherr, Yigong Shi, Bruce Stillman, Michael Stratton, George Thomas, Thea Tlsty, Eileen White, Junying Yuan, Qimin Zhan

Francis Crick Neuroscience Symposium

April 12 - 17, 2010
Abstract deadline: January 22, 2010

Arranged by:
Joshua Huang (USA)
Mu-ming Poo (China / USA)
Linda Richards (Australia)
Joshua Sanes (USA)
Keiji Tanaka (Japan)

Keynote Speaker
Linda Buck

Invited Speakers
Carol Barnes, Perry Bartlett, Tobias Bonhoeffer, Gyorgy Buzsaki, Ann-Shyn Chiang, Yang Dan, Karl Deisseroth, Shumin Duan, Catherine Dulac, Michael Fee, William Harris, Nobutaka Hirokawa, Christine Holt, Nancy Ip, Lily Jan, Yuh-Nung Jan, Gilles Laurent, Jeff Lichtman, Nikos Logothetis, Andreas Luethi, Liqun Luo, Minmin Luo, Jeffrey Macklis, Roberto Malinow, Edvard Moser, Fujio Murakami, Akinao Nose, Hitoshi Sakano, Erin Schuman, Shubha Tole, Li-Huei Tsai, Richard Tsien, Tony Zador

Topics:
- Neurogenesis (embryonic and adult)
- Neuronal development (polarization, axon guidance, migration)
- Synaptogenesis (formation, refinement)
- Synaptic plasticity cell biology of neurons (ion channels, trafficking, regulation)
- Neural circuit functions (sensory processing)
- Neural circuit functions (others)
- Cognitive functions (learning and memory)
- Cognitive functions (others)

Epigenetics, Chromatin & Transcription

May 17 - 21, 2010
Abstract deadline: February 26, 2010

Arranged by:
Rob Martienssen (USA)
Ru-Ming Xu (China)
Yi Zhang (USA)

Keynote Speakers
Thomas Jenuwein
Tony Kouzarides
Danny Reinberg
Robert Roeder

Topics:
- Histone and DNA modifications
- Nucleosome remodeling and non-coding RNAs
- Long-range chromatin interaction and chromatin domains
- Epigenetic modifications in development, differentiation, and diseases
- Global approaches to chromatin and transcription
- Structural and functional relationship for epigenetic factors and chromatin
- New techniques and new epigenetic phenomenons

Invited Speakers
Genevieve Almouzni, Sung Hee Back, Shelley Berger, Bradley Cairns, Xuemei Chen, Xiadong Cheng, Gerald Crabtree, Susan Gasser, Shiv Grewal, Peter Jones, En Li, Haifin Lin, Karolin Luger, Huck Hui Ng, Gang Pei, Wolf Reik, Michael Rosenfeld, Hiroyuki Sasaki, Eric Selker, Yang Shi, Ramin Shiekhattar, Kevin Struhl, Azim Surani, Teruhiko Wakayama, Jerry Workman, Richard Young, Kenneth Zaret, Jian-Kang Zhu

Please check our website for more details on upcoming symposia, conferences, workshops and summer schools in the life sciences and related disciplines. Programs will be held at the new Suzhou Dushu Lake Conference Center, sixty miles west of Shanghai and in easy reach of international and domestic airports. Suzhou is an ancient city known as the Venice of China, famous for its canals and private gardens and home to scholars for more than two millennia. Symposia, conferences, workshops and summer schools follow the Cold Spring Harbor tradition in showcasing merit-based scientific advances in an informal but intense atmosphere.

To encourage significant international participation from throughout Asia, Europe and North America, the language of all programs will be English.

Suzhou Dushu Lake Conference Center, SIP/Suzhou, Jiangsu Province, China Tel: 86/0 512 6272 9029 Fax: 86/0 512 6272 9028 www.csh-asia.org meetings@csh-asia.org

CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF MEDICINE

Full/Associate/Assistant Professor Case Western Reserve University School of Medicine

Case Center for Proteomics and Bioinformatics and Department of Molecular Biology and Microbiology

The Center for Proteomics and Bioinformatics and Department of Molecular Biology and Microbiology in the School of Medicine at Case Western Reserve University invite applications for an open tenure-track faculty position at the Full, Associate or Assistant Professor level. The successful candidate will establish a vigorous research program, participate in teaching activities, and interact productively with the strong team of basic and clinical scientists interested in the overall areas of HIV/AIDS and host-pathogen interactions, and drug abuse in the context of HIV/AIDS at our institution. We are especially interested in candidates who are undertaking epigenetic, proteomic, systems biology, and/or bioinformatics studies of HIV/AIDS. In addition to newly refurbished laboratory space and generous start-up packages, we offer a highly interactive environment with exceptional intellectual, infrastructural, and administrative support. The candidate will be an active participant in the expansion of HIV/AIDS research conducted by the Center for Proteomics and Bioinformatics and the Center for AIDS Research. The successful candidate will receive a primary appointment in either the Center for Proteomics and Bioinformatics or the Department of Molecular Biology and Microbiology and a secondary appointment in the collaborating department, depending on the candidate's research focus. All candidates should have a Ph.D. and relevant post-doctoral experience. Candidates for the rank of associate professor or above must have a record of funding, active research program and a national reputation. Rank commensurate with experience. Interested candidates should apply by **February 1, 2010** by sending a cv, a plan of research, and the names of 3 references to: **Joan Schenkel, Case Center for Proteomics and Bioinformatics, CWRU, 10900 Euclid Ave, Cleveland, OH 44106-4988** or **joan.schenkel@case.edu**.

In employment as in education, Case Western Reserve University is committed to Equal Opportunity and World Class Diversity. Case is a recipient of a National Science Foundation ADVANCE Institutional Transformation Grant to increase the participation of women in Science and Engineering.

CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF MEDICINE

Associate Professor
Case Center for Proteomics and Bioinformatics

The Center for Proteomics and Bioinformatics in the School of Medicine at Case Western Reserve University invites applications for a tenure track faculty position at the Associate Professor level. The areas of research interest are any areas in sequence based or structural bioinformatics for nucleic acids or proteins. Candidates interested in translational approaches, such as biomarker discovery or drug design, are also encouraged to apply. The candidate will also have the opportunity to collaborate with other Medical School faculty in a large number of NIH funded center and program project grants affiliated with the Center. These funded programs emphasize research in protein-protein interaction networks, integration of genomics and proteomics data, and high-throughput structure modeling, using a variety of methods. Overall, the candidate will be an active participant in shaping the expansion of bioinformatics and/or macromolecular structure programs of the Center. Candidates should have a Ph.D. and relevant academic or industrial experience suitable for appointment at an Associate Professor level.

Interested candidates should apply by **February 1, 2010** by sending a cv including funded grants, a plan of research and teaching, and the names of 3 references to: **Joan Schenkel, Case Center for Proteomics and Bioinformatics, CWRU, 10900 Euclid Ave, Cleveland, OH 44106-4988** or **joan.schenkel@case.edu**.

In employment as in education, Case Western Reserve University is committed to Equal Opportunity and World Class Diversity. Case is a recipient of a National Science Foundation ADVANCE Institutional Transformation Grant to increase the participation of women in Science and Engineering.

Faculty Positions

NUS National University of Singapore

The **Department of Biochemistry** of the National University of Singapore is embarking on a major expansion of its research program. We invite applications for full-time tenure-track Assistant Professors and tenured Associate Professor/ Full Professors. We are aiming at the frontier of biomedical sciences with an emphasis on biochemical mechanisms in neuronal systems including neuronal stem cells, vertebrate development, cytokine signaling, epigenetic control of learning and memory, therapeutic applications in major diseases and computational or systems biology. However, outstanding applicants in other areas are welcome to apply also.

National University of Singapore is a world-renowned research and educational institution. The Department has a strong research tradition in its 82 year history and active collaborations with top research organizations both in Singapore and internationals. The Department has excellent facilities for research in biochemistry and molecular and cell biology and success rates for funding of excellent research projects and programs are high. Personal compensation is highly competitive to those found in North America and Europe.

All faculty members are expected to teach undergraduate and/or graduate courses (although new appointments have a very light teaching load to allow them to establish their research programs), supervise graduate students, and conduct vigorous research programs that generate external funding and scholarship and intellectual output typical of that of a world-class university. Interested parties should submit their applications, supported by a resume, detailed research plan and names of three or more external referees to:

Professor Xin-Yuan Fu,
HEAD Department of Biochemistry
Yong Loo Lin School of Medicine
National University of Singapore
8 Medical Drive, MD7 #02- 03
Singapore 117597
Fax: (+ 65) 6 779 8842
E-mail: bchhead@nus.edu.sg
"http://www.med.nus.edu.sg/bioweb/

Only shortlisted candidates will be notified.

POSITIONS OPEN

The Department of Biological Sciences invites applications for the **ROBERT GRIGGS PROFESSOR,** a tenure-track **ASSISTANT** or **ASSOCIATE PROFESSORSHIP** in phylogenetic biogeography/coevolution with an appointment date starting in September 2010. We seek a **PHYLOGENETICIST** who uses comparative data to study historical biogeography or coevolution (for example, coevolution of parasites and their hosts or herbivores and their host plants), and whose teaching and research interests will expand and strengthen our Weintraub Program in Systematics and Evolution (see **website: http://www.gwu.edu/~clade**). The successful candidate will be expected to develop an externally funded research program and to participate in graduate and undergraduate education. Basic qualifications: A completed Ph.D. in a relevant field is required and postdoctoral experience is preferred. Applicants must demonstrate a commitment to scholarly research through publications, and they must have research experience using comparative data as described above. To be considered at the Associate Professor level, the candidate must already be at that rank or have seven years or more of experience as a faculty member or museum curator.

Review of applications begins December 16, 2009, and will be ongoing until the position is filled. Application procedure: Interested candidates should e-mail a letter of application, curriculum vitae, a statement of research and teaching interests, electronic copies of up to three publications, and have three letters of recommendation sent (Assistant Professor applicants) or provide a list of references (Associate Professor applicants) to:

Weintraub Search Committee Chair
Department of Biological Sciences
340 Lisner Hall, 2023 G Street N.W.
The George Washington University
Washington, DC 20052
E-mail: wsearch@gwu.edu
Telephone: 202-994-6090

Only complete applications will be considered.

The George Washington University is an Equal Opportunity/Affirmative Action Employer. The University Search Committee seeks to attract an active, culturally and academically diverse faculty of the highest caliber.

PRIZE FELLOWSHIPS to study Foundational Questions in Evolutionary Biology

The Program for Evolutionary Dynamics at Harvard University proposes to elect up to 10 **RESEARCH FELLOWS** to study foundational questions in evolutionary biology. The successful candidates will receive a grant of up to $75,000 per year for up to two consecutive years. The Fellows will do cutting-edge research at Harvard University or other academic institutions in the Boston area. The Fellows will be expected to participate in the creation of new research networks, to attend regular meetings, and to work across disciplinary boundaries. Work which has significant philosophical implications for evolutionary biology, and science more broadly, will receive priority. Applicants must have a Ph.D. and an excellent track record of scientific productivity, as evidenced by publications. Applications including a brief description of research interests and proposed projects (no more than three pages) and curriculum vitae should be sent electronically to **e-mail: lydia_liu@harvard.edu** by February 1, 2010. Three letters of reference should also be directly submitted to the same address. One letter should come from the principal investigator of the research laboratory where the applicant plans to work. Further information is available at **website: http://www.ped.fas.harvard.edu.**

A **POSTDOCTORAL POSITION** is available to study protein function at Fox Chase Cancer Center. The position is funded by NIH. We study protein-protein interactions that regulate the macromolecular assembly and intracellular trafficking of membrane receptors and ion channels. The position requires a Ph.D. in cell biology or membrane protein biochemistry/biophysics. Highly motivated candidates who can work independently are encouraged to apply. **E-mail: zimei.bu@fccc.edu.**

POSITIONS OPEN

FACULTY POSITIONS
IN ENERGY SCIENCE OR ENGINEERING
Stanford University

Stanford University, in conjunction with its Precourt Institute for Energy, seeks nominations and applications for two faculty appointments in energy science and/or engineering. The Precourt Institute (**website: http://pie.stanford.edu**) is Stanford's interdisciplinary hub for energy research and teaching. These are tenure-track positions that will be joint appointments in the Institute and a department appropriate to the candidate in the School of Earth Sciences, the School of Engineering, or the School of Humanities & Sciences. The appointments are expected to be at the **ASSISTANT PROFESSOR** or **ASSOCIATE PROFESSOR** (without tenure) level.

Possible areas of expertise include, but are not limited to, the following: conversions of energy from a renewable resource (wind, solar, geothermal, among others), energy conversions that reduce greenhouse gas emissions, energy systems that enable greater use of renewable resources (an improved electric power grid, for example), sustainable energy, materials science of energy conversions, biofuels, energy storage on various time scales, and electrochemistry and catalysis of energy conversions.

The successful candidate is expected to conduct a vigorous research program and to be an active participant in the Precourt Institute for Energy (**website: http://pie.stanford.edu**) interdisciplinary energy initiatives campuswide, and in teaching and mentoring students, in accordance with particular interests and expertise. Examples of such initiatives are the Precourt Energy Efficiency Center (**website: http://piee.stanford.edu**), the Global Climate and Energy Project (**website: http://gcep.stanford.edu**), and the Stanford Institute for Materials and Energy Science (**website: http://www-public.slac.stanford.edu/simes/**). Additional departments and programs engaged in energy research and teaching at Stanford are listed at **website: http://pie.stanford.edu/assets.html**. Candidates who currently hold faculty positions are expected to have a demonstrated record of excellence in research and teaching. Junior candidates should have completed a Doctorate before the date of appointment and show evidence of outstanding potential for research and teaching.

Applicants are required to provide the following materials: a cover letter describing research and teaching experience as well as future plans in these areas, curriculum vitae, copies/reprints of up to five publications, and a list of three to five references with complete contact information. The committee will request letters of recommendation for finalists. Please electronically submit the requested materials in PDF format to **e-mail: fmorr@stanford.edu**. Applicants who are unable to submit the materials in electronic form may mail the requested materials to:

Professor Franklin M. Orr, Jr.
Search Committee Chair
Yang & Yamazaki Environment & Energy Building
Suite 324, 473 Via Ortega, MC 4230
Stanford, CA 94305

Review of applications will begin on 15 December 2009 and will continue until the positions are filled.

Stanford University is an Equal Opportunity Employer and is committed to increasing the diversity of its faculty. It welcomes nominations of and applications from women and minority groups, as well as others who would bring additional dimensions to the University's research, teaching, and clinical missions.

POSITIONS OPEN

FACULTY POSITIONS
Yale University School of Medicine
Translational Neuroscience in the Department of Neurology

The Yale Department of Neurology is initiating searches for as many as five **NEUROSCIENTISTS** who use molecular, cellular, genetic, electrophysiological, and/or imaging approaches to advance the understanding of pathophysiology or treatment of neurological diseases. We will consider applicants in any area of neuroscience, including neurodegeneration and repair, neuro-oncology, neuroimmunology/virology, and neurogenetics. Priority will be given to applicants with research programs of direct clinical relevance, and to physician-scientists with clinical expertise in addition to scientific accomplishments.

Candidates must hold an M.D. and/or a Ph.D. or equivalent degrees. We invite applications at the rank of **ASSISTANT PROFESSOR,** but appointments at the rank of **ASSOCIATE PROFESSOR** or **PROFESSOR** will be considered. Applications are due by December 30, 2009. Please send a cover letter, curriculum vitae, up to three representative publications, and a research plan (strictly limited to two pages) and arrange for submission of three letters of recommendation.

All application materials should be sent electronically to **David A. Hafler** or **Stephen M. Strittmatter,** exclusively through the following address, **e-mail: sheila.macmillan@yale.edu.**

Applications from women and minority scientists are especially encouraged. Yale is an Affirmative Action/Equal Opportunity Employer.

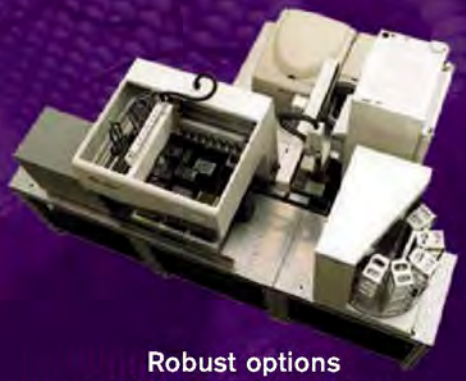
PROTEOMICS OF AGING. Applications are invited for National Institute of Standards and Technology/National Research Council **POSTDOCTORAL** program (**website: http://sites.nationalacademies.org/pga/rap/**). *The program is restricted to U.S. citizens* who have held the Doctorate less than five years. Areas of research include mass spectrometry of membrane proteins. The experience in protein chemistry or mass spectrometry would be an advantage. Starting salary will be $63,400 plus fringe benefits. The deadline for final application is February 1, 2010. Prior to formal application, prospective applicants should contact: **Dr. Illarion V. Turko, Center for Advanced Research in Biotechnology, Rockville, MD 20850. E-mail: turko@umbi.umd.edu.**

MARKETPLACE

Achieve more with cellular automation

Versatile configurations for many cellular processes.

Compact workcells that simplify and automate cellular assays



Robust options and devices for process cell culture configurations

Automate your cell culture and assay process from start to finish.

Automate antibody addition for flow cytometry preparation:

- Easy user interface
- Add antibodies to plates or tubes

Create a workstation ideal for higher throughput cell-based assays:

- Data input to modify assay during run
- Schedule assays with exact timings and error prediction

Monitor process cell culture over days or months:

- Test pH, lactate, glucose & viability
- Many detector options
- Automated centrifugation

Choose software options tailored for your process:

- Assay Optimization software (AAO) for improved assay robustness
- Capture all the details of an assay run in an easy Excel interface

Contact us to create a solution that maximizes your cellular discovery process!

Visit www.beckmancoulter.com/LabAutoSolutions

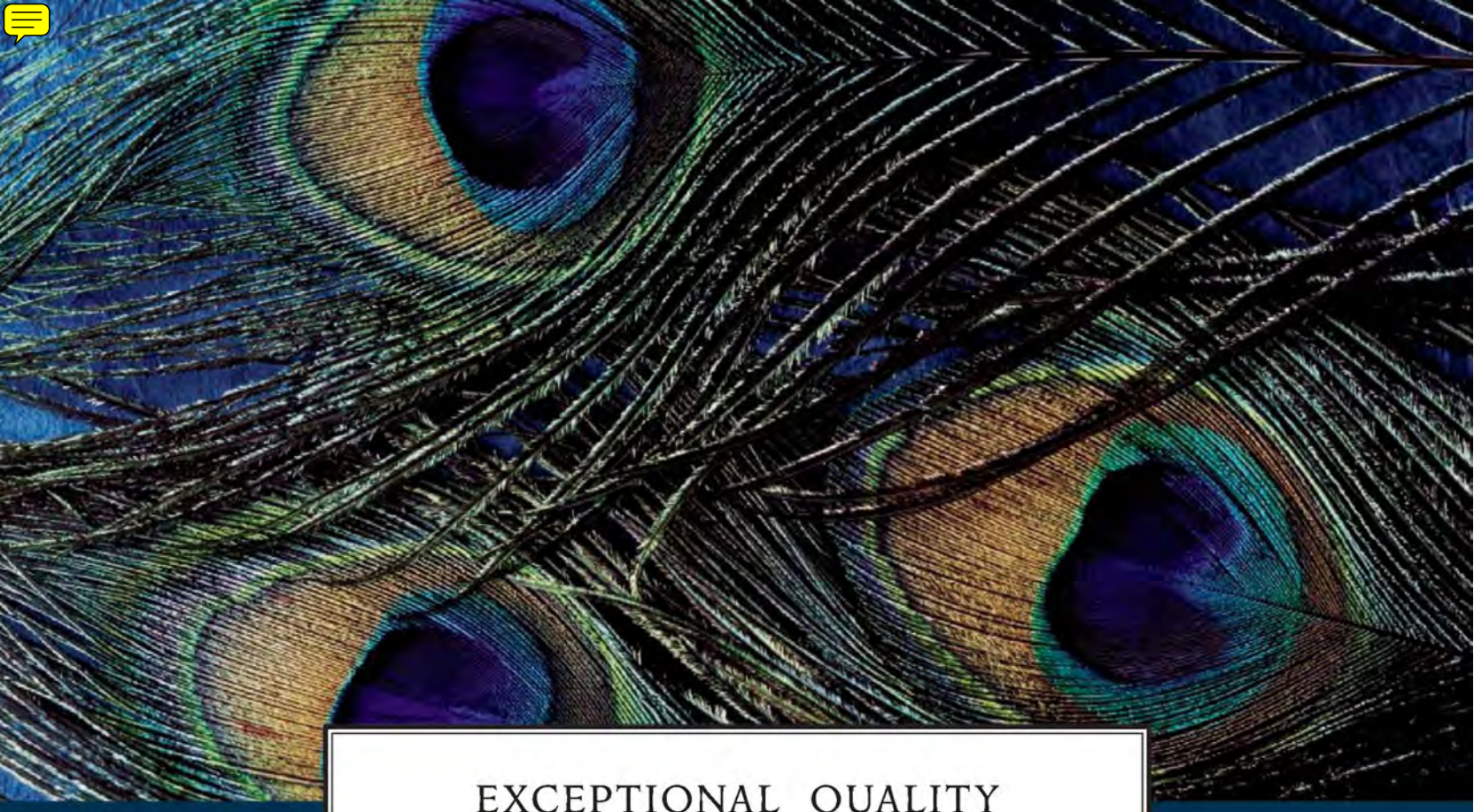
Beckman Coulter Automation systems are part of the **2009 Stimulus PLUS Grant Award Program,*** which offers special savings on Beckman Coulter products.

Visit www.beckmancoulter.com/Stimulus

*Valid in the US only

Genomics Cell Analysis Particle Characterization Capillary Electrophoresis **Lab Automation** Centrifugation Bioseparation Lab Tools

BECKMAN COULTER
We're better together

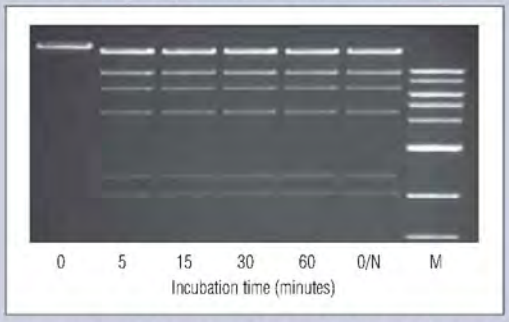


EXCEPTIONAL QUALITY

Restriction Enzymes from New England Biolabs

With 35 years of experience in enzyme technology, New England Biolabs leads the industry in quality, selection and innovation for restriction enzymes. Our expertise in purification coupled with a reliance on recombinant technology results in highly pure enzymes suitable for even the most sensitive applications. Additionally, stringent quality controls ensure that our enzymes deliver optimal performance. For quality you can count on, make NEB restriction enzymes your first choice.

Power and purity of NEB Restriction Enzymes



Recombinant HindIII is powerful enough to digest 1 µg of DNA in 5 minutes, but can also be used in overnight digests with no indication of nuclease contamination. Marker (M) is the 1 kb DNA Ladder (NEB #N3232).

BB *Choose recombinant enzymes for a higher quality product with lot-to-lot consistency*

Advantages:

- **Selection** – More specificities than any other supplier
- **Convenience** – Optimal activity for over 160 enzymes in a single buffer NEB4
- **Quality** – State-of-the-art production and stringent QCs
- **Innovation** – HF enzymes engineered for reduced star activity
- **Performance** – Guaranteed

NEW ENGLAND BioLabs® Inc.
enabling technologies in the life sciences

CLONING & MAPPING	DNA AMPLIFICATION & PCR	RNA ANALYSIS	PROTEIN EXPRESSION & ANALYSIS	GENE EXPRESSION & CELLULAR ANALYSIS

www.neb.com